UNIVERSITY OF LIVERPOOL

Investigating Polo-like kinase 4-regulated signalling using SILAC-based quantitative phosphoproteomics

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

Samantha Ferries

September 2018

Acknowledgements

Thank you to Professor Rob Beynon and everyone in the Centre for Proteome Research. I have really enjoyed working with you all and I am so grateful for the support, friendship and guidance you have all given me throughout my PhD. I would like to thank Dr. Stephen Holman for his excellent teaching of the principles of mass spectrometry.

Thank you to Dr. David Mason for the immunofluorescence experiments and to Professor. Sonia Rocha for helping me with my FACs experiments. Also, thank you to Dr. Simon Perkins for helping me (a lot!) with the bioinformatics. I am very grateful for all your help.

I would like to give a special thank you to both Dr. Dominic Byrne and Dr. Philip Brownridge. Dom, thank you for everything. Thank you for training me to such a high standard in so many methods and for always being there to give endless advice and help. And Philip, thank you for everything you have taught me about mass spectrometry. It has been a privilege to work with and learn from you.

A huge thank you to Becka, Grace and Gemma for the nights out, the endless cups of tea and our cheese and wine nights. I could not have got through my PhD without you. And a huge thank you to Amy. I am so glad we could go through this process together and thank you for your help with everything.

To my friends Amy, Beth, Jess and Kate, thank you so much for always being there for me. I am so lucky to have you as my friends and your support the past four years has meant everything to me.

To my Mum, Dad, Laura, James, and of course George; thank you for always believing in me and encouraging me to achieve my goals. I would not be where I am today without your support.

Finally, to Phil, thank you for your support and for helping me believe in myself. Your support, particularly at the end of my PhD, was invaluable and I couldn't have done it without you.

Abstract

Title: Investigating PLK4- regulated signalling pathways using SILAC-based quantitative phosphoproteomics

Author: Samantha Ferries

Protein phosphorylation plays a critical role in regulating cellular responses, including the co-ordination of cell division. Polo-like kinase 4 (PLK4), a cell cycle regulated Ser/Thr protein kinase, is the master regulator of centriole biogenesis, a process required for formation of the biopolar mitotic spindle. PLK4 expression and activity are both very tightly controlled, and dysregulation leads to aberrant centrosome duplication resulting in chromosome missegregation and aneuploidy. Despite its rising importance as a potential anti-cancer target, few PLK4 substrates have been identified, and much remains to be understood about basic PLK4 biology at the centrosome, and wider potential roles within the cell. The experiments described in this thesis focused on an evaluation of PLK4 signalling, exploiting SILAC-based quantitative phosphoproteomics and the protein kinase inhibitor centrinone to probe cellular function.

Initially, I describe the development of biochemical tools to study PLK4-dependent signalling, and its inhibition by centrinone. This included the generation of stable isogenic U2OS cell lines, transfected with either WT PLK4 or a centrinone-resistant G95R PLK4 mutant. The G95R PLK4 cell line was designed to explore potential off-target effects of centrinone inhibition, and biochemical and cellular assays confirmed the 'drug-resistant' status of this protein. An optimal concentration of centrinone was established to ensure inhibition of WT PLK4, whilst minimising off-target effects.

Next, mass spectrometric (MS) acquisition parameters for phosphopeptide analysis were optimised on an Orbitrap Fusion, using a synthetic phosphopeptide library and phosphopeptide-enriched cell lysate. Eight acquisition methods were assessed, considering mode of fragmentation (HCD, EThcD, and neutral loss triggered ET(ca/hc)D), and analysers for MS2 (orbitrap and ion trap) in addition to assessment of optimal downstream processing. HCD, with orbitrap MS2 analysis provided high numbers of identifications with high phosphosite localisation and was used to study the PLK4 regulated phosphoproteome.

Building upon this, phosphoproteomics subsequently revealed 412 significantly ($p \leq 0.075$) regulated proteins in WT PLK4 expressing cells following centrinone treatment, and 471 regulated phosphosites, with 313 downregulated (inhibited) and 158 upregulated (activated) after compound exposure. A target of centrinone was implicated in MAPK signalling and G1/S phases of the cell cycle. In addition, regulation of biological processes including rRNA processing, cell-adhesion and transcription was observed. A novel PLK4 proline directed consensus phosphorylation motif was also identified from phosphoproteomic data and validated biochemically with recombinant PLK4 *in vitro*.

Finally, PLK4 was demonstrated to be a biochemical upstream activator of Aurora A kinase, phosphorylating it at Thr288 within the activation loop, which led to hyperactivation. In addition, both proteins co-localised in human cultured interphase cells, suggesting potential novel functions for PLK4-regulated Aurora A in the early stages of the cell cycle. Taken together, this work provides a series of novel insights into PLK4-regulated signalling, which will be useful for future studies evaluating the biological and disease-related functions of this important fundamental protein kinase.

Table of contents

List of Figures

List of Tables

Abbreviations

adj	Adjusted
ADP	Adenosine diphosphate
AGC	Automatic gain control
AI-NETD	Activated ion-negative electron transfer dissociation
amu	Arbitrary mass units
ANOVA	Analysis of variance
aPK	Atypical protein kinase
bp	Base pairs
BP	Biological Process
CC	Cellular compartment
CDK	Cyclin-dependent kinase
CID	Collision-induced dissociation
DDA	Data-dependent acquisition
DIA	Data-independent acquisition
DSF	Differential scanning fluorimetry
ECL	Enhanced chemiluminescence
ePK	Conventional protein kinase
ESI	Electrospray ionisation
ETcaD	Electron transfer collisional activated dissociation
ETD	Electron transfer dissociation
EThcD	Electron transfer higher-energy collisional dissociation
FDR	False discovery rate
FLR	False localisation rate
FRT	FLP recombination target
GO	Gene ontology
H/L	Heavy to light
HCD	Higher-energy collisional dissociation
IC_{50}	Half maximal inhibitory concentration
IT	Ion trap
iTRAQ	Isobaric tags for relative an absolute quantitation
KD	Kinase domain
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
m/z	Mass-to-charge ratio

MF	Molecular function
MS	Mass spectrometry
MW	Molecular weight
NCE	Normalised collision energy
NL	Neutral loss
OT	Orbitrap
PBD	Polo box domain
PD	Proteome Discoverer
PLK	Polo-like kinase
PSM	Peptide spectrum match
PTM	Post-translational modification
R0K0	Control media supplemented with light arginine and lysine
R10K8	Media supplemented with heavy $^{13}C_6$ $^{15}N_4$ arginine and $^{13}C_6$ $^{15}N_2$ lysine
RP-HPLC	Reversed phase high-pressure liquid chromatography
SD	Standard Deviation
SILAC	Stable-isotope labelling by amino acids in cell culture
TCL	Total cell lysates
Tet	Tetracycline
U2OS	Human bone osteosarcoma epithelial cells
WT	Wild type

Amino Acids

Amino acid	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

pSer, pThr, pTyr = phosphorylated amino acids

Publications

Below is a list of publications I have authored (1) or co-authored (6) during my PhD:

Foulkes, D. M., Byrne, D. P., Yeung, W., Shrestha, S., Bailey, F. P., Ferries, S., Eyers, C. E., Keeshan, K., Wells, C., Drewry, D. H., Zeurcher, W. J., Kannan, N. & Eyers, P. A (2018). Covalent EGFR/HER2 kinase inhibitors induce cellular degradation of human Tribbles 2 (TRIB2) pseudokinase. Sci. Signal. Just accepted.

Eyers, C. E., Vonderach, M., Ferries, S., Jeacock, K. & Eyers, P. A. (2018). 'Understanding protein-drug interactions using ion mobility-mass spectrometry'. Curr Opin Chem Biol. 42. 167-176.

Bury, L., Coelho, P. A., Simeone, A., Ferries, S., Eyers, C. E., Eyers, P. A., Zernicka-Goetz, M. & Glover, D. M. (2017). 'Plk4 and Aurora A cooperate in the initiation of acentriolar spindle assembly in mammalian oocytes'. J Cell Biol. 216. 3571- 3590.

Ferries, S., Perkins, S., Brownridge, P. J., Campbell, A., Eyers, P. A., Jones, A. R. & Eyers, C. E. (2017). 'Evaluation of Parameters for Confident Phosphorylation Site Localization Using an Orbitrap Fusion Tribrid Mass Spectrometer'. J Proteome Res. 16. 3448-3459.

McSkimming, D. I., Dastgheib, S., Baffi, T. R., Byrne, D. P., Ferries, S., Scott, S. T., Newton, A. C., Eyers, C. E., Kochut, K. J., Eyers, P. A., Kannan, N. (2016). 'KinView: a visual comparative sequence analysis tool for integrated kinome research'. Mol Biosyst. 12. 3651-3665.

Byrne, D. P., Vonderach, N., Ferries, S., Brownridge, P. J., Eyers, C. E. & Eyers, P. A. (2016) 'cAMP-dependent protein kinase (PKA) complexes probed by complementary differential scanning fluorimetry and ion mobility-mass spectrometry'. Biochem J. 473. 3159-3175.

Mohanty, S., Oruganty, L., Kwon, A., Byrne, D. P., Ferries, S., Ruan Z., Hanold, L. E., Katiyar, S., Kennedy, E. J., Eyers, P. A. & Kannan, N. (2016) 'Hydrophobic Core Variations Provide a Structural Framework for Tyrosine Kinase Evolution and Functional Specialization', PLoS Genet, 12, doi: 10.1371/journal.pgen.1005885

1. Introduction

1.1. Protein Kinases

Eukaryotic cells are continuously exposed to extracellular signals, all of which must be transduced and interpreted downstream, leading to specific cellular responses. Intracellular signalling is directed by post-translational modifications, of which phosphorylation is one of the most prevalent (Hunter, 1995). Protein phosphorylation is driven by the activity of protein kinases, which reversibly add phosphate groups via the hydrolysis of ATP to a number of amino acids, of which serine, threonine and tyrosine are the most widely studied (Figure 1.1). There is also some evidence that phosphorylation can occur on residues including histidine, arginine, lysine, aspartate, glutamate and cysteine (Hardman *et al.*, 2017). Phosphate groups on Ser, Thr and Tyr are removed by the subsequent regulated activity of protein phosphatases. Phosphorylation at specific sites within a protein leads to a diverse range of 'downstream' effects, including protein activation, deactivation, regulating of protein stability and subcellular localisation. All of these effects are integrated to drive cellular processes including cell division, differentiation and apoptosis, amongst many others (Manning *et al.*, 2002; Bononi *et al.*, 2011).

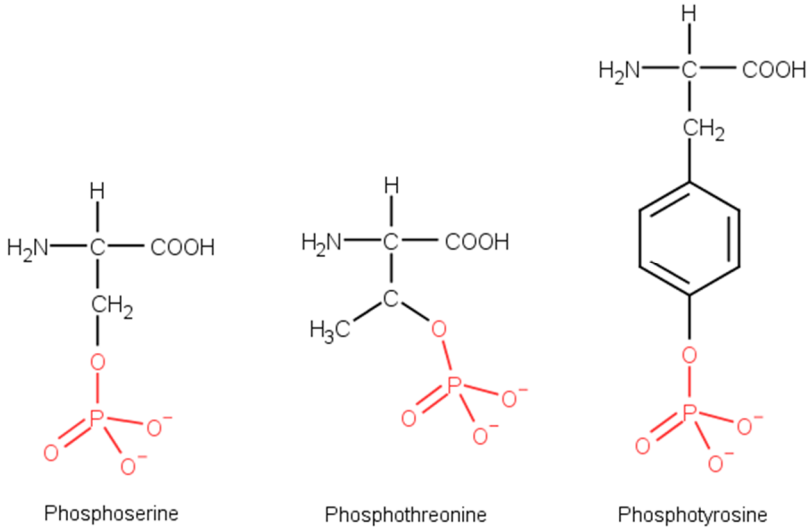


Figure 1.1 Structure of phosphoserine, phosphothreonine and phosphotyrosine

Over 500 protein kinases are encoded in the human genome and have been classified in to a hierarchy of groups depending on sequence similarity, overall domain structure, substrate specificity and mode of regulation (Hanks & Hunter, 1995). Conventional protein kinases (ePKs) are classified in to eight groups (Figure 1.2), whereas atypical protein kinases (aPKs), defined as those without sufficient sequence similarity to be classified as an ePK, are divided in to four separate groups.

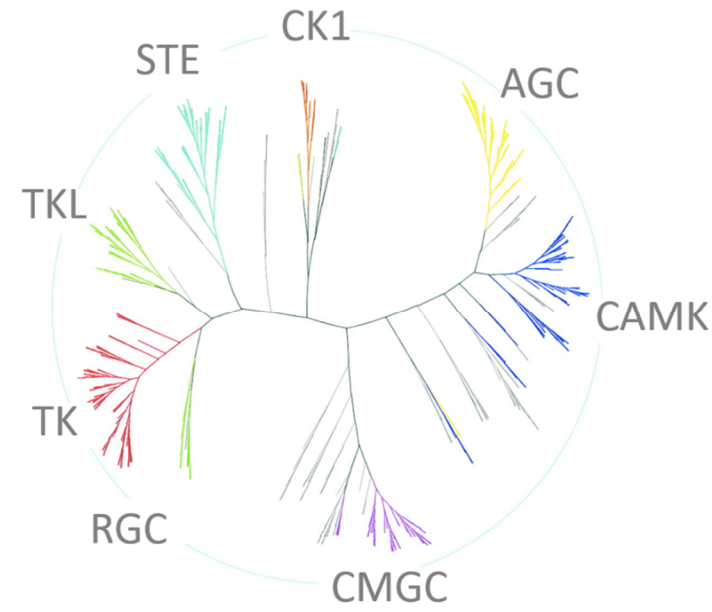


Figure 1.2 Classification of conventional protein kinases. Conventional protein kinases are divided in to 8 groups depending on sequence similarity, domain structure, substrate specificity and mode of regulation. Figure adapted from Manning *et al*., 2002.

Protein kinases incorporate a number of unique structural features including a conserved core of two lobes (Knighton *et al.*, 1991). The N-lobe consists of a five stranded β-sheet and an αC helix, whereas the C-lobe is mostly helical (Figure 1.3. A). A cleft between the two lobes forms the active site, which binds ATP co-ordinated to two divalent cations (magnesium or manganese) (Kornev *et al.*, 2006).

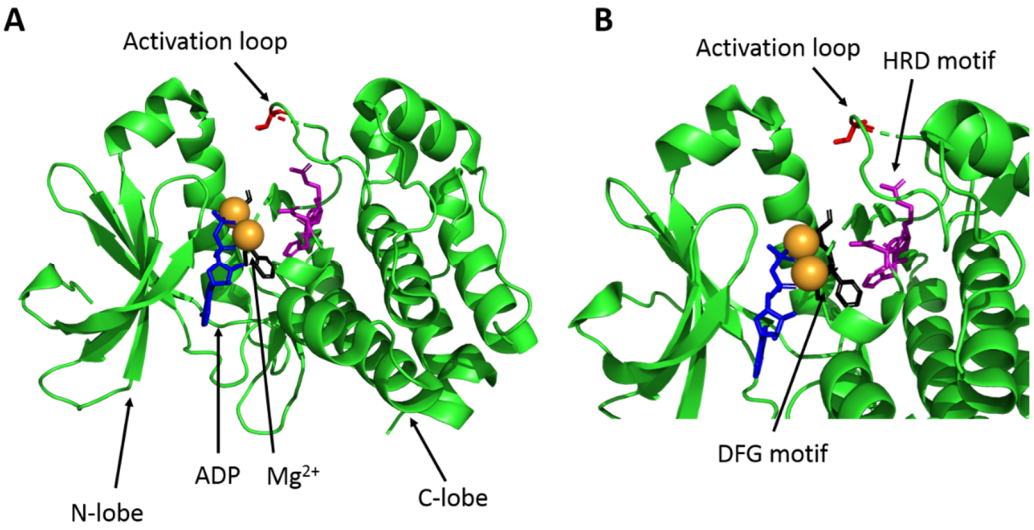


Figure 1.3 Domain structure of protein kinases. A) Tertiary structure of the Ser/Thr protein kinase, Aurora A, highlighting the N- and C- lobes, the activation loop and the bound ADP and Mg^{2+}. B) The DFG and HRD regulatory motifs positioned in Aurora A. Images were produced in PyMOL.

Protein kinases contain a number of conserved regulatory features, including the activation segment. The activation segment is phosphorylated as the kinase becomes active and forms an extended conformation to allow substrate binding. The activation loop is stabilised by an invariant aspartate located within a DxWxxG motif in the C-lobe and the APE motif, which lies at the C-terminal end of the activation segment. An invariant glutamate residue in the N-lobe forms a salt bridge with a key lysine residue that is essential for protein kinase activity (Figure 1.4) (Schenk & Snaar-Jagalska, 1999; Spitaler *et al.*, 2000). The activation segment is positioned adjacent to the Asp-Phe-Gly (DFG) motif, a key regulatory motif that adopts different conformations depending on whether the kinase is in its active or inactive state. In active kinases, the DFG motif plays a key role in catalysis and adopts an 'in' conformation that orientates the motif toward the bound ATP (Huse & Kuriyan, 2002). The Asp residue is then correctly positioned to coordinate to the divalent cations bound to the β- and γ- phosphate groups of ATP in the active site (Figure 1.3. B; Figure 1.4). Another important regulatory feature of protein kinases is the His-Arg-Asp (HRD) motif, which is positioned correctly upon phosphorylation of the activation loop (Figure 1.3. B; Figure 1.4). In the active kinase, the Asp residue of this motif acts as a catalytic base for nucleophilic attack of the hydroxyl group of the substrate required for transfer of the γ-phosphate (Hanks *et al.*, 1988; Johnson *et al.*, 1996). A glycine rich loop at the N-terminus is a key part of the ATP binding site, which anchors and shields the bound ATP (Hemmer *et al.*, 1997). Also within the ATP binding site is the important 'gatekeeper residue' which confers

selectivity for nucleotide binding, and is frequently mutated upon targeting of a kinase with an inhibitor to confer drug resistance (Figure 1.4) (Bishop *et al.*, 2000; Treiber & Shah, 2013). Highly controlled phosphorylation of protein substrates by protein kinases is a direct result of the dynamic regulation of these key structural features within the kinase active site.

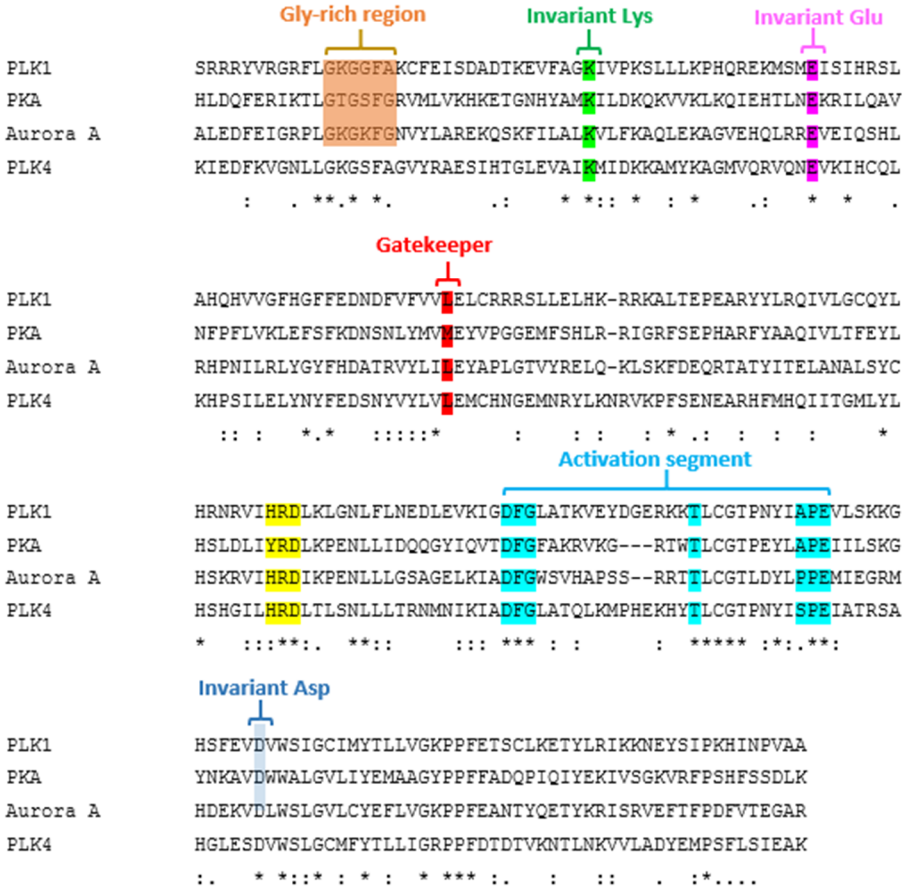


Figure 1.4 Conserved protein kinase structural features. The sequence alignments of PLK1, PKA, Aurora A and PLK4 demonstrate conserved structural elements including the Gly-rich ATP binding loop (orange); invariant lysine residue, essential for kinase activity (green); invariant glutamate which forms a salt bridge with the invariant lysine (purple); the gatekeeper residue responsible for nucleotide and inhibitor binding selectivity (red); the HRD motif, required for catalysis (yellow); the activation segment, incorporating the DFG motif required for co-ordination of the bound divalent cations and APE motif which stabilises the activation segment (light blue), along with the invariant aspartate (dark blue).

1.2. The cell cycle

Protein kinases play a major role in regulating the cell cycle. For a somatic cell to undergo mitosis, it must first replicate its organelles and DNA before partitioning the chromosomes into the daughter cell. The cell cycle is divided in to distinct stages: interphase, prophase, prometaphase, metaphase, anaphase and telophase. Each stage of the cell cycle is carefully controlled by the co-ordinated activities of protein kinases to ensure the fidelity of the process (Schafer, 1998).

Interphase describes the processes required to prepare the cell for mitosis and can be divided in to G1, S and G2 (Figure 1.5). Cyclin-dependent kinases (CDKs) bound to specific cyclins are the major regulators of cell division and act to phosphorylate a number of substrates to drive proliferation, while other families of protein kinases (including Polo-like and Aurora kinases) regulate specific mitotic events (Nigg, 2001; Malumbres & Barbacid, 2005).

In G1, the cells grow larger and duplicate their organelles and must pass a 'restriction point', defined as the stage at which the cell has committed to dividing and no longer responds to growth factors. CDK4-cyclin D is responsible for co-ordinating the activities in G1 and in comparison to other CDK-cyclin complexes, requires additional cofactors to become active (Baldin *et al.*, 1993; Kato *et al.*, 1994); the cofactor Cip/Kip proteins (p21, p27 and p57) play both an activating and inhibitory role depending on their phosphorylation state (Cheng *et al.*, 1999). Unphosphorylated Cip/Kip proteins inhibit CDK4-cyclin D activity, whereas phosphorylation of specific tyrosine residues results in a conformational change, preventing inhibition and releasing CDK4-cyclin D to phosphorylate downstream substrates (James *et al.*, 2008).

Once active, CDK4-cyclin D regulates the restriction point via its interaction with retinoblastoma protein (RB1). Hypophosphorylated RB1 interacts with E2F transcription factors, sequestering them to prevent transcription of downstream targets (Ikeda *et al.*, 1996). CDK4-cyclin D phosphorylation of RB1 on specific residues releases E2F, which transcribes genes involved in DNA synthesis, marking the transition from G1 to S phase (Zarkowska & Mittnacht, 1997; Rubin *et al.*, 2005).

At the G1/S phase transition, levels of another CDK-cyclin complex, CDK2-cyclin E increase (Lauper *et al.*, 1998). This complex, in conjunction with Cdc6, stimulates the assembly of the pre-replication complex, preparing the chromatin for replication (Lunn *et al.*, 2010). Once this has been achieved, cyclin A levels increase and form complexes with CDK2 to initiate DNA synthesis (Coverley *et al.*, 2002). Cells then progress into G2 and continue to grow in preparation for entry to mitosis. Cyclin A is degraded, whilst cyclin B is synthesised and forms a complex with CDK1. CDK1 becomes active following dephosphorylation at Thr14 and Tyr15 by Cdc25 phosphatases (Sebastian *et al.*, 1993). CDK1-Cyclin B then phosphorylate a wide range of substrates including lamins, microtubule binding proteins and condensins that regulate important M phase processes such as breakdown of the nuclear envelope and spindle assembly (Blethrow *et al.*, 2008).

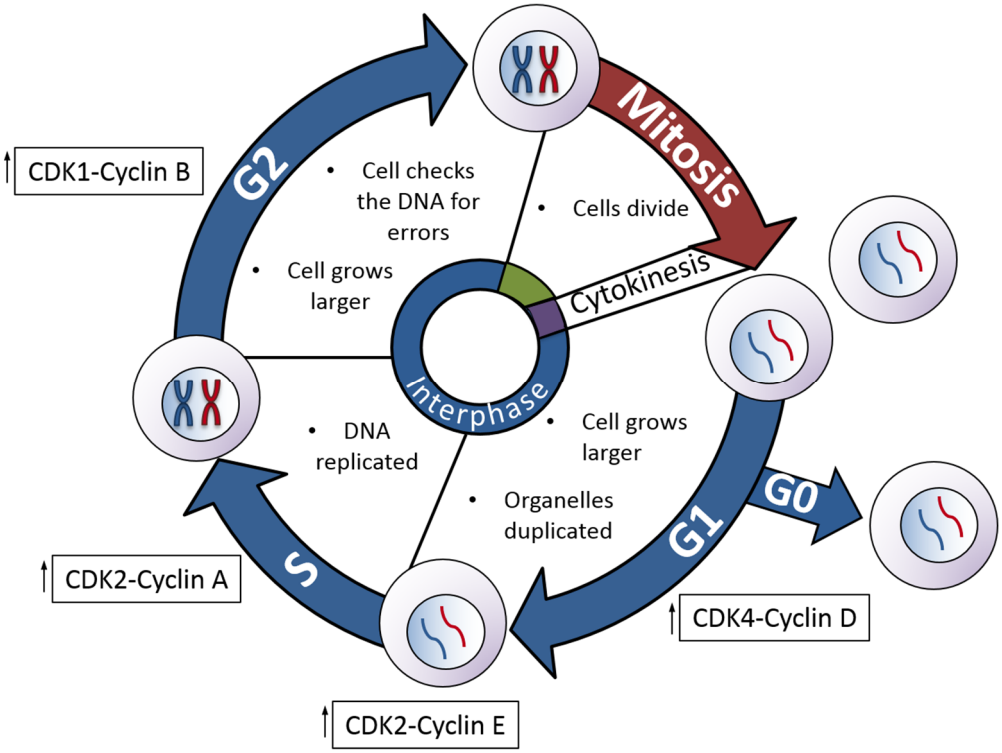


Figure 1.5 Stages of interphase in the eukaryotic cell cycle. Interphase is divided in to G1, S and G2 phase, preparing the cell for entry in to mitosis.

In prophase, the chromatin begins to condense via the action of condensin complexes, whilst cohesin complexes keep sister chromatids together. In addition, the nuclear envelope begins to break down to release the DNA so that the chromosomes can associate with spindle fibres as mitosis progresses (Salina *et al.*, 2002). The centrosomes begin to migrate to opposite poles of the cell and start to assemble spindle fibres. In prometaphase,

the spindle fibres continue to assemble and chromosomes develop kinetochores which are required to attach sister chromatids to the spindle (Figure 1.6) (Kirschner & Mitchison, 1986).

In metaphase, the centrosomes have migrated to opposite poles of the cell and chromosomes are accurately aligned along the metaphase plate (Rieder *et al.*, 1986). In anaphase, cohesin complexes are cleaved by the protease separin and the separated chromatids are pulled towards opposite poles of the cell (Waizenegger *et al.*, 2000). The chromosomes start to decondense in telophase, and the nuclear envelope reforms to surround the DNA. A contractile ring then forms under the plasma membrane and constricts, forming a cleavage furrow that splits the daughter cells in two (Glotzer, 2005). The cells then re-enter interphase and prepare for another round of cell division.

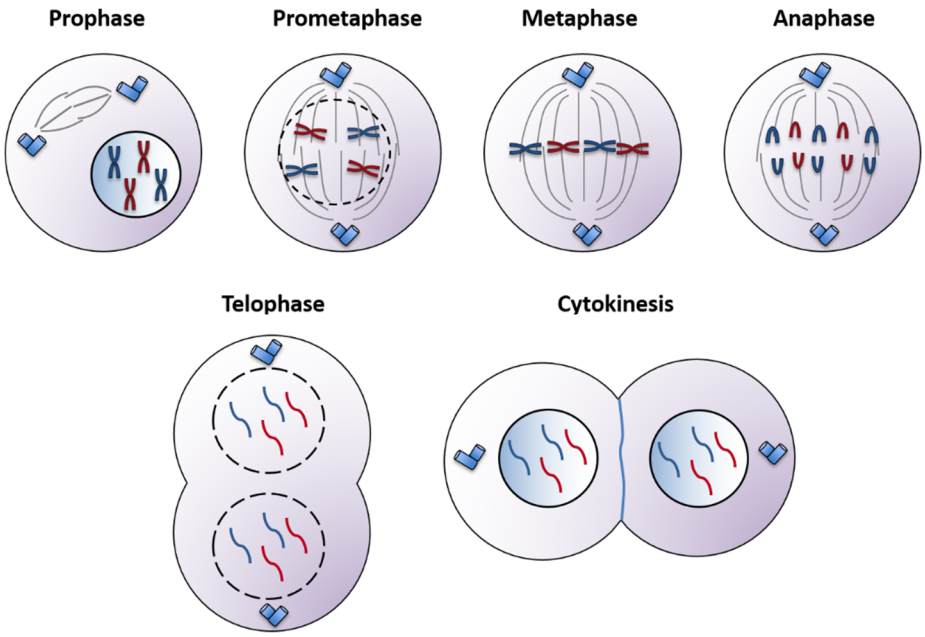


Figure 1.6 Stages of mitosis in the eukaryotic cell cycle. Mitosis is divided in to six stages: prophase, prometaphase, metaphase, anaphase, telophase and cytokinesis.

1.3. The centrosome cycle

Centrosomes are the microtubule organising centres of the cell. They are composed of two microtubule-based structures called centrioles, which are formed with nine-fold symmetry and are surrounded pericentriolar material (PCM) containing a matrix of proteins. (Kochanski & Borisy, 1990; Dammermann *et al.*, 2004). The centrosome is responsible for nucleating microtubules to form the spindle fibres required for chromosome segregation during the cell cycle. In addition, the centrosome also regulates cilia formation by functioning as basal bodies and thus plays an important role in cell signalling (Goetz & Anderson, 2010).

Each cell contains one centrosome and is duplicated only once per cell cycle in a highly controlled process at the G1/S transition. Exactly one procentriole forms perpendicular to each of the two parent centrioles, which continue to elongate, mature and are finally released from the parent centriole in late mitosis (Figure 1.7). This process is regulated by a number of key proteins, including members of the Polo-like kinase family and Aurora kinases. The role of these kinases in co-ordinating activities at the centrosome is described in detail below.

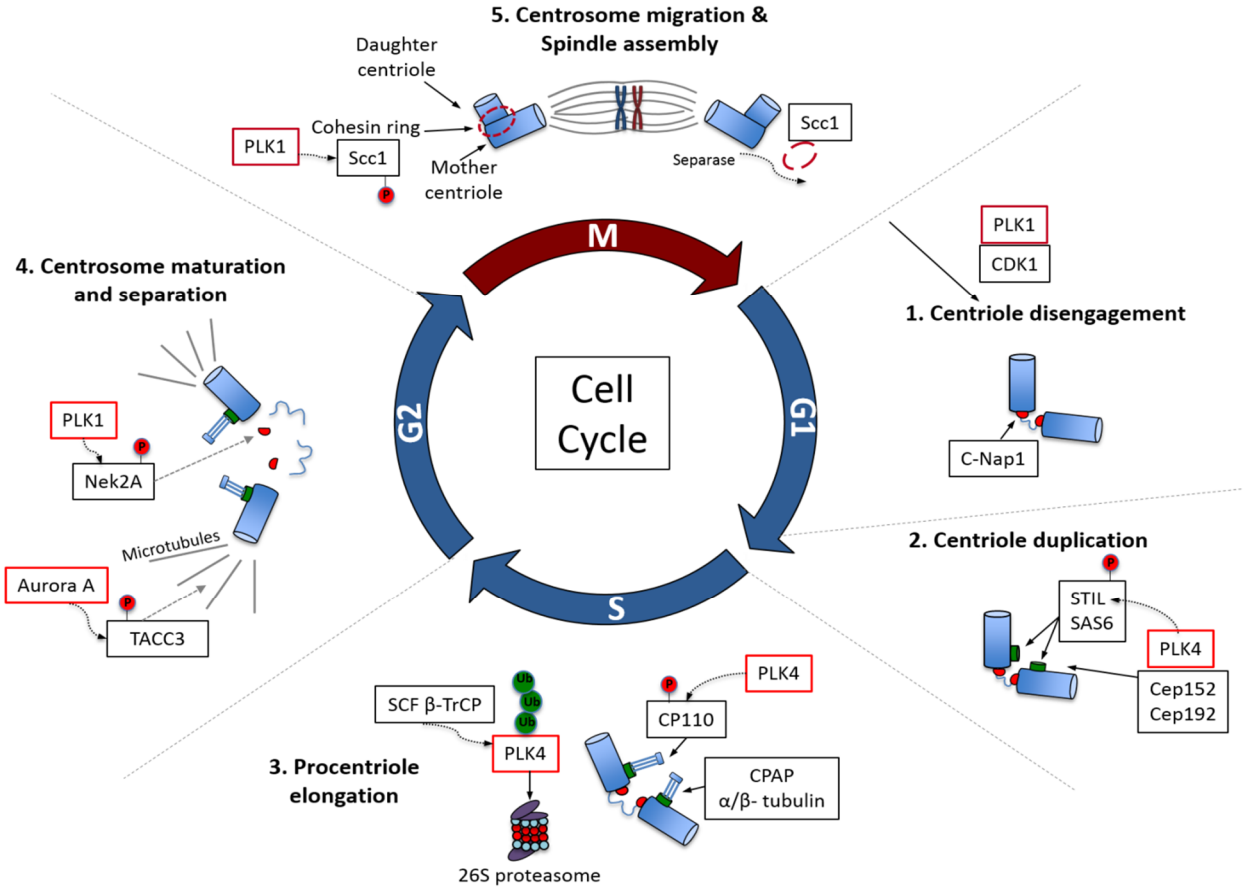


Figure 1.7 The centrosome cycle. The centrosome cycle begins with centriole disengagement (1) prior to the recruitment of PLK4, STIL and SAS6 to initiate centriole biogenesis at the G1/S. transition. In S phase, the procentriole elongates and PLK4 is degraded by the 26S proteasome (3). In G2, the centrosome undergoes maturation via Aurora A phosphorylation of substrates including TACC3 and the centrosome separates by the activity of PLK1 phosphorylated Nek2A (4). In mitosis, the centrioles migrate and assemble the spindle fibres. The cohesin ring holding the mother and daughter centrioles together is degraded to complete the cycle (5). Adapted from Wang *et al*., 2014.

1.4. Polo-like kinases

The family of Polo-like protein kinases were first discovered in *Drosophila,* when mutant alleles of the *polo* locus were identified, which resulted in the development of abnormal spindle poles (Sunkel & Glover, 1988). Five members of the Polo-like kinase (PLK) family have been identified (Figure 1.8). PLK1-3 share a similar architecture, with an N-terminal Ser/Thr kinase domain and two C-terminal regulatory domains termed polo box domains (PBD) responsible for subcellular localisation and substrate specificity (Lowery *et al.*, 2005). PLK4 is structurally divergent, as it contains three PBDs rather than two (Leung *et al.*, 2002). A fifth member, PLK5 is identified by the absence of a functional kinase domain (de Carcer *et al.*, 2011).

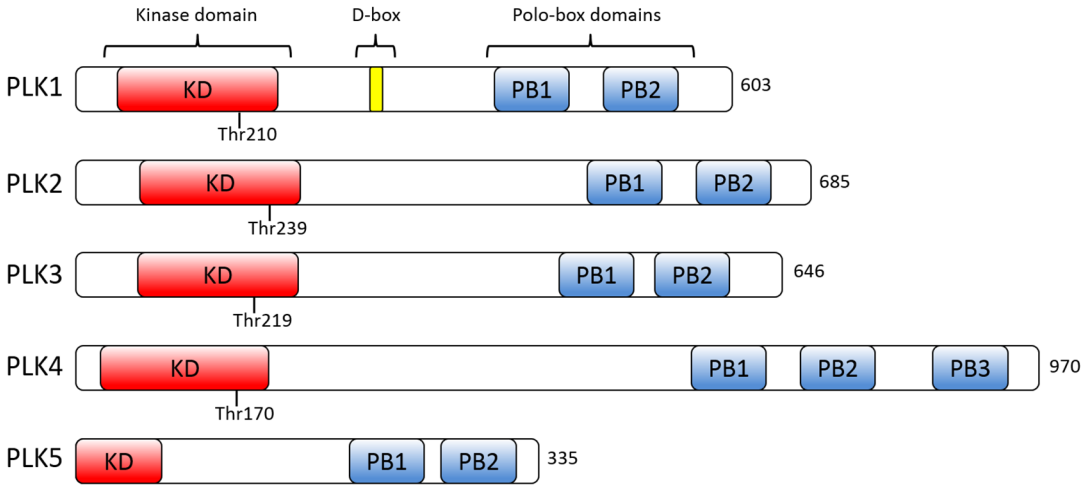


Figure 1.8 Schematic of the domain structures in Polo-like kinases. The domain structure of polo-like kinases is illustrated, depicting the N-terminal kinase domain (KD), and C-terminal polo-box domains (PB) involved in regulation. The destruction box (D-box) of PLK1 is shown in yellow. The number of amino acids of each kinase is shown on the right. The labelled Thr residues represent the activating residue within the kinase activation segment.

1.4.1. **PLK1**

PLK1 is the most widely studied polo-kinase and has important roles across the cell cycle, including at the centrosome. At the G2/M phase transition, PLK1 is involved in centrosome maturation via phosphorylation of pericentrin. This recruits key proteins to the centrosome including γ-tubulin, GCP-WD and the scaffold protein Cep192, involved in centrosome maturation (Lee & Rhee, 2011). PLK1 also functions to separate the centrosomes so that spindle assembly can occur prior to segregation of the chromosomes. In interphase, centrosomes are joined together by proteins including C-Nap1 and rootletin (Yang *et al.*, 2006). These proteins are displaced upon entry to mitosis by PLK1 phosphorylation of Nek2A kinase, which acts at the centrosome linker allowing the centrosomes to separate (Mardin *et al.*, 2011) (Figure 1.7).

As the cell exits mitosis, PLK1 aids in separation of the mother and daughter centrioles by phosphorylating a cehesin ring subunit, Scc1, which can then be cleaved by separase, releasing the centrioles (Losada *et al.*, 2002; Hauf *et al.*, 2005) (Figure 1.7). In addition to these roles in centrosome biology, PLK1 has extensive functions across the cell cycle, including mitotic entry in conjunction with CDK1-cyclin B, spindle assembly, chromosome segregation and cytokinesis (Jackman *et al.*, 2003; Burkard *et al.*, 2007).

PLK1 is overexpressed in many cancers and widely considered as an oncogene. NIH 323 cells constitutively expressing PLK1 formed tumours when injected into nude mice, suggesting that PLK1 contributes to the progression of human cancers (Smith *et al.*, 1997). In prostate cancer cells, PLK1 overexpression is linked to inactivation of PTEN, and loss of PTEN is believed to drive the majority of prostate cancers (Liu *et al.*, 2011). However, recent work has revealed that dysregulation of PLK1 can reduce aberrant cell proliferation in cancer cells. Overexpression in mammary tumour cells led to abnormal chromosome segregation and a failure of the cell to undergo cytokinesis, suppressing tumour development (de Carcer *et al.*, 2018). PLK1 may therefore possess both oncogenic and tumour suppressive functions, depending on the cancer under study.

1.4.2. **PLK2**

PLK2 localises to the centrosome and is activated at the G1/S transition. PLK2 is involved in centriole duplication via a destabilising phosphorylation of three conserved residues in FBXW7 (a substrate recognition component of an SCF E3 ubiquitin ligase). FBXW7 regulates the degradation of target proteins, including cyclin E and therefore, a decrease in FBXW7 leads to an accumulation of cyclin E in the cell. In addition to the roles described above, cyclin E, bound to CDK2, localizes to the centrosomes where the complex can phosphorylate centrosomal substrates including nucleophosmin (NPM) thereby promoting centriole duplication (Okuda *et al.*, 2000; Cizmecioglu *et al.*, 2012).

PLK2 has typically been considered as a tumour suppressor, as it is transcriptionally regulated by p53 (Burns *et al.*, 2003). In addition, silencing of PLK2 expression occurs in malignant B cells, highlighting its tumour suppressive effects (Syed *et al.*, 2006). However, overexpression of PLK2 has been identified in colorectal cancer cells where hyperactivity and uncontrolled degradation of FBXW7 leads to accumulation of FBXW7 target proteins including cyclin E, driving uncontrolled cell growth in this cancer type (Ou *et al.*, 2016). It is therefore possible that PLK2, like PLK1, is both a tumour suppressor and oncogene depending on cell type.

1.4.3. **PLK3**

PLK3 is expressed throughout the cell cycle and is activated during G2. It is found in the cytoplasm, at the cell membrane and in the nucleus. Similar to the other PLKs, PLK3 has roles in the cell cycle, including the indirect control of cyclin E levels. PLK3 phosphorylates the protein phosphatase Cdc25a, which acts to remove inhibitory phosphate groups from cyclin E dependent kinases, promoting entry into S phase (Blomberg & Hoffmann, 1999; Zimmerman & Erikson, 2007).

PLK3 also has roles outside of the cell cycle including cellular responses to stress. Oxidative stress activates p53, which regulates transcription of genes involved in the stress response. This activation is, in part, mediated by PLK3 via phosphorylation at Ser20 of p53, leading to cell cycle arrest and apoptosis. PLK3 activity therefore provides a link between the DNA damage response and apoptosis (Xie *et al.*, 2001).

The cellular responses to oxidative stress suggest that PLK3 functions as a tumour suppressor. PLK3 has been shown to negatively regulate HIF-1α, a master transcriptional regulator of genes in response to hypoxia. HIF-1α activity enables survival of cancer cells in hypoxic conditions and therefore, its negative regulation by PLK3 highlights the tumour suppressive functions of this kinase (Yang *et al.*, 2008; Xu *et al.*, 2010).

1.4.4. **PLK4**

PLK4 is a serine/threonine protein kinase, first discovered in mouse, with shared homology with the *Drosophila* polo kinase (Fode *et al.*, 1994). PLK4 localises at the centrosome and is the master regulator of centriole duplication, ensuring centrosomes are replicated only once per cell cycle (Habedanck *et al.*, 2005; Kleylein-Sohn *et al.*, 2007).

PLK4 is a low abundance protein and levels in the cell are tightly regulated to control centriole duplication and maintain genomic integrity. PLK4 is degraded via ubiquitin-mediated degradation by the 26S proteasome. The E3 ubiquitin ligase SKP1-CUL1-F Box (SCF) complex contains a β-TrCP F box protein that recognises its substrates upon phosphorylation in specific regions, promoting protein degradation. PLK4 forms homodimers and undergoes trans-autophosphorylation within a specific destruction (DSG) motif (residues 282-305), known as the phosphodegron. This is recognised by the E3 ubiquitin ligase which ubiquitylates PLK4 within its kinase domain to signal it for proteasomal degradation (Holland *et al.*, 2010; Sillibourne *et al.*, 2010; Cunha-Ferreira *et al.*, 2013). PLK4 is stabilised by the activities of protein phosphatase 2A, which dephosphorylates PLK4, preventing its autodestruction and allowing licenced centriole duplication to occur (Brownlee *et al.*, 2011).

An additional E3 ubiquitin ligase, Mib1, has been identified as a regulator of PLK4 levels in the cell. Mib1 has traditionally been associated with Notch and NF-κB signalling. However, Mib1 has also been shown to localise to centrosomes upon initiation of centriole biogenesis. At the centrosome, Mib1 ubiquitylates PLK4 at fifteen Lys residues across the kinase and polo-box domains and the Lys29- and Lys48 ubiquitin linkages formed resulted in proteasomal degradation. Mib1 ubiquitylation of PLK4 is not mediated by the SCF β-TrCP, and it is hypothesised that SCF β-TrCP controls PLK4 abundance in the cell cycle whereas Mib1 ubiquitylation acts on overexpressed PLK4 (Cajanek *et al.*, 2015).

Studies in *Drosophila* revealed that PLK4 also trans-autophosphorylates at Thr172 (Thr170 in humans) within its activation loop following accumulation at the centrioles (Lopes *et al.*, 2015). This results in autoactivation of PLK4 and permits the phosphorylation of its substrates to initiate centriole duplication. PLK4 is recruited to the centrosome by the scaffold proteins Cep152 and Cep192 and is capable of phosphorylating Cep152 in the process of preparing the molecular machinery required for centriole biogenesis (Figure 1.7) (Bahtz *et al.*, 2012; Kim *et al.*, 2013). Another key PLK4 substrate is GCP6, a member of the γ–tubulin ring complex (γ–TuRC). This complex localises to the PCM and distal ends of the centrosomes and has a role in nucleating microtubules. GCP6 has been shown to be required for centriole duplication and it is regulated by phosphorylation within a tandem repeat segment by PLK4 (Bahtz *et al.*, 2012).

To initiate centriole duplication, PLK4 phosphorylates another of its substrates: the centrosomal localised protein STIL. STIL is recruited to the centrosome by the centrosomal scaffold protein Cep85, which binds STIL in its N-terminal domain (Liu *et al.*, 2018). STIL protein undergoes self-oligomerization via its conserved CC domain, which is also required to bind PLK4 (Arquint *et al.*, 2015). Once bound, PLK4 activity is stimulated and phosphorylates STIL at a number of residues within its STAN domain, including Ser1108 and Ser1116 (Moyer *et al.*, 2015). Phosphorylated STIL then recruits SAS6, a structural protein involved in forming the nine-fold symmetry typical of microtubules, forming a new centriole (Dzhindzhev *et al.*, 2017) (Figure 1.7; Figure 1.9). PLK4 also phosphorylates the coiled-coil protein CP110, which acts antagonistically with the centrosomal protein CPAP (also recruited by Cep152) to determine centriole length. Phosphorylated CP110 acts to stabilise SAS6 during cartwheel formation to establish the nine-fold symmetry of the growing centriole (Lee *et al.*, 2017). SAS-6 levels in the cell are regulated by the SCF-FBXW5 E3 ubiquitin ligase, which adds ubiquitin groups to SAS6 to signal the protein for degradation following the formation of the nascent centriole. FBXW5 activity must be controlled to prevent premature ubiquitylation of SAS6 and it has been shown that FBXW5 is negatively regulated by PLK4 phosphorylation at Ser151. Following centriole duplication, PLK4 autophosphorylation and subsequent degradation relieves the SCF-FBXW5 complex to ubiquitylate SAS6 and prevent additional centrosome duplication events from occurring (Puklowski *et al.*, 2011).

An additional mechanism to prevent multiple duplication events involves CDK1-cyclin B activity in the later phases of the cell cycle. A portion of PLK4, STIL and SAS6 can be identified at the mother and daughter centrioles following centrosome duplication (Leidel *et al.*, 2005; Arquint *et al.*, 2012; Kratz *et al.*, 2015) however, CDK1-cyclin B binding to the CC domain of STIL prevents complex formation. CDK-cyclin B binds at the CC domain and phosphorylates STIL at the N-terminus. This prevents PLK4 binding to STIL and reduces PLK4 autoactivation, preventing overduplication of the centrosomes (Zitouni *et al.*, 2016). Furthermore, PLK4 activity is negatively regulated by acetylation of two Lys residues within the catalytic domain. The Lys acetyltransferase (KAT2A and KAT2B) acetylates PLK4 on Lys45 and Lys46, causing a conformational shift to the inactive state, inhibiting PLK4 activity and preventing further amplification of the centrosomes (Fournier *et al.*, 2016). Taken together, a number of important regulatory components act to control the levels of active PLK4 to ensure the fidelity of centriole duplication to once per cell cycle.

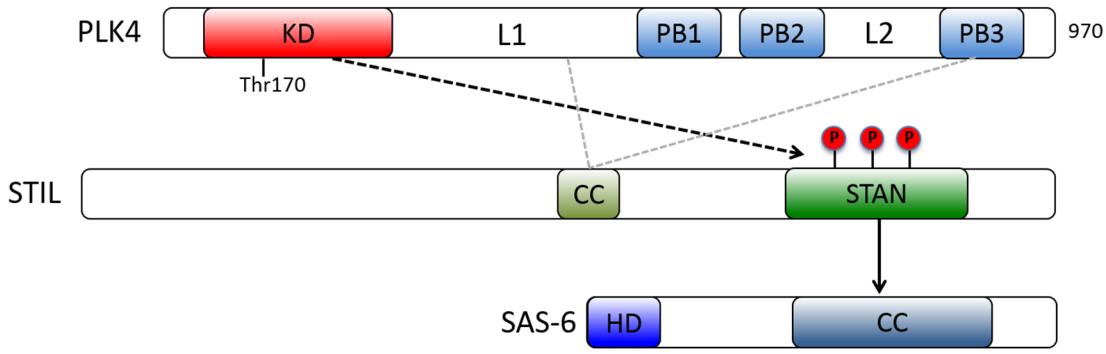


Figure 1.9 The PLK4-STIL-SAS-6 interaction. PLK4 binds to the CC domain of STIL via the L1 linker and polo-box 3. PLK4 then phosphorylates STIL within its STAN domain. Phosphorylated STIL then recruits SAS-6 to the centrosome to initiate centriole biogenesis. Adapted from Arquint & Nigg, 2016.

Overexpression of PLK4 leads to the production of supernumerary centrioles and over amplification of centrosomes. This leads to missegregation of chromosomes, which is a hallmark of many cancers (Habedanck *et al.*, 2005). A direct role in cancer progression was identified in *Drosophila*, where PLK4 overexpression in neuroblasts promoted tumorigenesis (Basto *et al.*, 2008). Mouse models (p53 null mice) have shown that PLK4 overexpression and centrosome amplification results in hyperproliferation and early onset tumour formation. In addition, overexpression resulted in a failure to produce basal bodies and primary cilia, which greatly disrupted cell signalling and caused uncontrolled

differentiation of the epidermis (Coelho *et al.*, 2015). In humans, PLK4 overexpression has been observed in many cancers, including breast, colorectal and gastric cancers (Macmillan *et al.*, 2001; Shinmura *et al.*, 2014; Li *et al.*, 2016). A role in metastasis was demonstrated via the interaction of PLK4 with the Arp2/3 complex responsible for regulating the actin cytoskeleton. PLK4 binding to Arp2/3 results in phosphorylation of the complex that is required for cytoskeletal rearrangement and subsequent cell motility. This suggested a mechanism of PLK4 driven cell metastases in human cancers (Kazazian *et al.*, 2017).

Loss of PLK4 results in a failure to duplicate the centrioles (Coelho *et al.*, 2015; Wong *et al.*, 2015). Without centriole duplication, the dividing cells will contain a single centrosome, resulting in the formation of aberrant mitotic spindles, causing errors in chromosome segregation. Loss of PLK4 has also been identified as a causal factor of microcephaly and retinopathy (Martin *et al.*, 2014; Tsutsumi *et al.*, 2016). Mutations in PLK4 transcripts were identified in patients with microcephaly, including a deletion mutation and premature truncation of PLK4 within PB3. This greatly reduced the amount of functional PLK4 in the cell and consequently impaired centriole duplication. Studies in zebrafish confirmed that mutated PLK4 led to a reduction in overall size and impaired vision, which are the known phenotypes of microcephaly and retinopathy (Martin *et al.*, 2014).

Taken together, dysregulation of PLK4, leading to either over- or under-expression, and consequently altered levels of activity, is implicated in a range of diseases and disorders. A comprehensive understanding of the global roles of PLK4 is therefore required to fully elucidate the importance of the involvement of this kinase in the cell cycle, at the centrosome and in other cellular processes. This is the main focus of the work presented in this PhD thesis.

1.4.5. **PLK5**

PLK5 lacks a functional kinase domain due to the presence of a stop codon disrupting the open reading frame. PLK5 does not appear to have roles in cell division and instead functions in neuron development. It is believed to serve a tumour suppressive function as it induces apoptosis in astrocytoma and glioblastoma tumour cells (de Carcer *et al.*, 2011).

1.5. Aurora kinases

The Aurora kinases were first discovered in *Drosophila* with homologs subsequently identified in all other eukaryotes (Glover *et al.*, 1995). There are three members of this kinase family: Aurora A, B and C, each of which share conserved structural features (Figure 1.10). Aurora kinases have a regulatory domain at the N-terminus and a C-terminal catalytic domain. Aurora A and B have an A box at the N-terminus, implicated in kinase degradation, which is missing in the shorter Aurora C. However, all three contain a D-box at the C-terminus, which is also responsible for regulating degradation. Despite the catalytic domains sharing near identical sequence identity, Aurora A has 50 times greater activity than Aurora B, thought to be due to a single key residue, G205 in Aurora A, which results in a conformational change bringing the αC-helix closer to the activation loop (Eyers *et al.*, 2005).

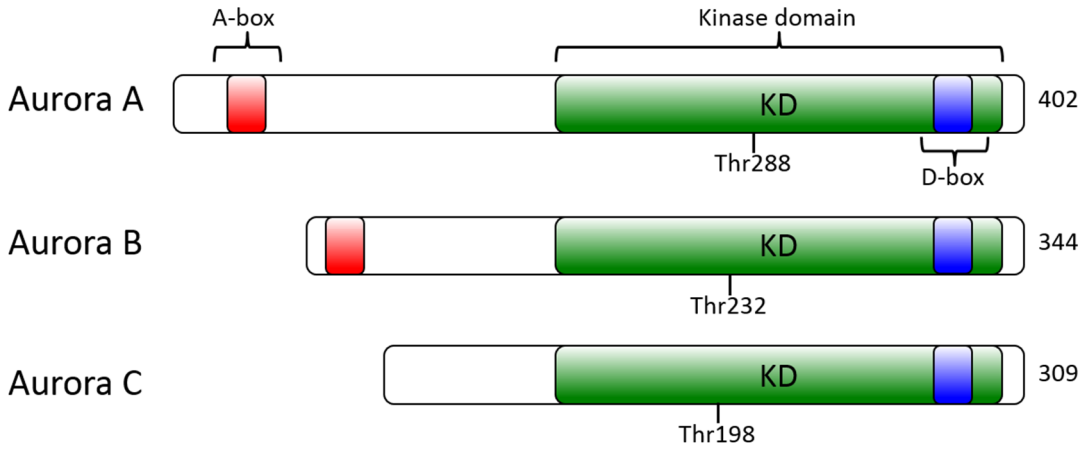


Figure 1.10. **Schematic of the domain structures in Aurora kinases**. The domain structure of Aurora kinases is illustrated, depicting the C-terminal kinase domain (KD) and the destruction box (D-box). A secondary destruction box, the A-box is found on Aurora A and Aurora B. The number of amino acids of each kinase is shown on the right. The labeled Thr residues represent the activating residue within the kinase activation segment.

1.5.1. **Aurora A**

Aurora A has a diverse range of functions including roles at the centrosome, entry in to mitosis, and assembly of the spindle fibres (Hannak *et al.*, 2001; Marumoto *et al.*, 2002). Aurora A activity is regulated by phosphorylation at Thr288 within its activation segment (Walter *et al.*, 2000) and its activity toward protein substrates is facilitated by the binding of cofactors including TPX2, Ajuba and Bora (Eyers *et al.*, 2003; Hirota *et al.*, 2003; Hutterer *et al.*, 2006).

In G2 phase cells, Aurora A is recruited to the centrosome by upstream regulators CDK11 and PAK1 (Zhao *et al.*, 2005; Petretti *et al.*, 2006). At the centrosomes, Aurora A is activated by the LIM protein Ajuba. Active Aurora A then recruits proteins involved in centrosome maturation, including γ-tubulin and the TACC3-Msps complex. Aurora A activates this complex which then stabilizes the microtubules, a key aspect of centrosome maturation (Figure 1.7) (Barros *et al.*, 2005). In addition, Aurora A phosphorylates Cdc25b which becomes active and dephosphorylates CDK1-cyclin B at Thr14 and Tyr15 residues (Cazales *et al.*, 2005). Activated CDK1-cyclin B then drives the cell to enter mitosis.

TPX2 was identified as an activator of Aurora A required for it to function in spindle assembly (Eyers *et al.*, 2003). The N-terminus of TPX2 binds to the C-terminus of Aurora A, inducing conformational changes which both enhance autophosphorylation at Thr288 within the activation loop, and shield the phosphorylated residue from the activity of downstream phosphatases (Bayliss *et al.*, 2003). TPX2 bound Aurora A has been identified in a complex with microtubule proteins including XMAP215 and HURP, which act together to drive spindle assembly (Wong *et al.*, 2008).

Aurora A degradation is mediated by the anaphase-promoting complex/cyclosome (APC/C). This ubiquitin ligase targets a wide range of substrates, resulting in proteolytic degradation of mitotic proteins as the cell prepares to exit mitosis. The APC/C, bound to a co-activator Cdh1, recognises the degradation motifs (D-box and A-box) within Aurora A (Taguchi *et al.*, 2002). Then, two APC/C associated enzymes; Ube2C and Ube2S add ubiquitin groups leading to rapid degradation of the kinase.

Overexpressed Aurora A has been identified in a range of cancers, including breast, colorectal, ovarian and pancreatic (Bischoff *et al.*, 1998). Aberrant Aurora A results in a

failure of the cell to undergo cytokinesis and centrosome amplification via multinucleation of microtubules (Meraldi *et al.*, 2002). In addition, cells overexpressing Aurora A exhibit reduced p53 activity. This is due to phosphorylation at specific residues by Aurora A, preventing DNA binding and transcription of downstream target genes. This phosphorylation also enhances interaction between p53 and Mdm2 protein, promoting p53 degradation (Liu *et al.*, 2004b). Taken together, Aurora A expression must be carefully controlled to ensure the fidelity of cell cycle processes.

1.5.2. **Aurora B**

Aurora B functions in the later stages of the cell cycle and its roles include chromatin modification, attachment of the spindle microtubules to kinetochores and cytokinesis. Localisation and activation of Aurora B is dependent on the regulatory subunits of the chromosome passenger complex (CPC): INCENP, Survivin and Borealin (Bolton *et al.*, 2002). Binding of INCENP to Aurora B results in allosteric activation via a conformational change of the activation loop. In a feedback mechanism, Aurora B phosphorylates INCENP to become fully active (Honda *et al.*, 2003). At centromeres, Borealin concentrates Aurora B leading to increased autoactivation.

Aurora B bound to the CPC regulatory subunits plays a key role in regulating kinetochore attachments to the spindle, to ensure accurate segregation of sister chromatids. Phosphorylation of kinetochore proteins, such as NDC80 and KNL1, result in a destabilising effect between kinetochore-microtubule interactions. This is required to ensure that the chromosomes can re-align and eventually bi-orient correctly for segregation (Cheeseman *et al.*, 2006; Welburn *et al.*, 2010).

CPC bound Aurora B also plays a key role in cytokinesis (Giet & Glover, 2001). In anaphase, Aurora B localises to the central spindle midzone and phosphorylates proteins involved in cytokinesis including vimentin present at the cleavage furrow, myosin II and desmin. Aurora B ensures the correct localisation and activity of the centralspindlin complex. This complex is required by the cytokinesis regulator RhoA, which signals the formation of the contractile ring to separate the cytoplasm in to the daughter cells (Basant *et al.*, 2015).

Dysregulation of Aurora B has been identified in gliomas, thyroid carcinomas and colon cancers. Overexpression disrupts Aurora B activities at the kinetochore, preventing the accurate segregation of chromosomes, resulting in aneuploid cells (Gonzalez-Loyola *et al.*, 2015).

1.5.3. **Aurora C**

Aurora C is the least studied of the Aurora family. It has been shown to localise to the centrosomes in a similar pattern to Aurora B. It can also interact with the CPC subunits INCENP and Survivin (Sasai *et al.*, 2016). Aurora C is mostly expressed in the testis and is required for spermatogenesis (Tseng *et al.*, 1998). Its role at the centrosome in somatic cells has not yet been studied. However, an increase in Aurora C mRNA levels have been observed in breast and prostate cancer cell lines, suggesting Aurora C may have a functional role in cancer progression (Zekri *et al.*, 2012).

1.6. Protein kinases as drug targets

Aberrant phosphorylation is implicated in all stages of neoplasia, including uncontrolled cell division and therefore, protein kinases have been identified as key chemotherapeutic targets. The first kinase inhibitor to reach the market was Imatinib, designed to treat chronic myelogenous leukaemia (CML) by inhibiting the proto-oncogene BCR-ABL (Deininger *et al.*, 1997). The success of this compound has led to the development of many small molecule inhibitors targeting the ATP binding site of protein kinases for the treatment of disease, including cancer. Protein kinase inhibitors are divided in to type I, type II and type III (Dar & Shokat, 2011). Type I inhibitors bind to the active, DFG 'in' conformation in an ATP competitive manner, and use the gatekeeper residues within kinases to gain selectivity. Type II inhibitors bind to the inactive, DFG 'out' conformation and type III inhibitors are non-ATP competitive and confer allosteric inhibition.

The highly conserved catalytic domain of protein kinases poses a vast challenge for the development of small-molecule compounds, particularly ATP competitive type I inhibitors. These inhibitors have multiple targets leading to the possibility of unknown off-target effects. Type II and type III inhibitors exhibit greater specificity toward their targets and overcome this issue to some degree. However, an additional issue arises from the ability of

protein kinases to mutate key residues to confer drug resistance. This was identified in patients treated with Imatinib, where a significant proportion developed mutations in the protein, including the gatekeeper residue (T315I) (Gorre *et al.*, 2001). This led to the development of next generation ABL inhibitors including Dasatinib, Nilotinib and Ponatinib (Figure 1.11) (Shah *et al.*, 2004; O'Hare *et al.*, 2009; Blay & von Mehren, 2011). The development of kinase inhibitors is continuously advancing in pursuit of developing new, highly specific inhibitors and overcoming the issues of drug resistance.

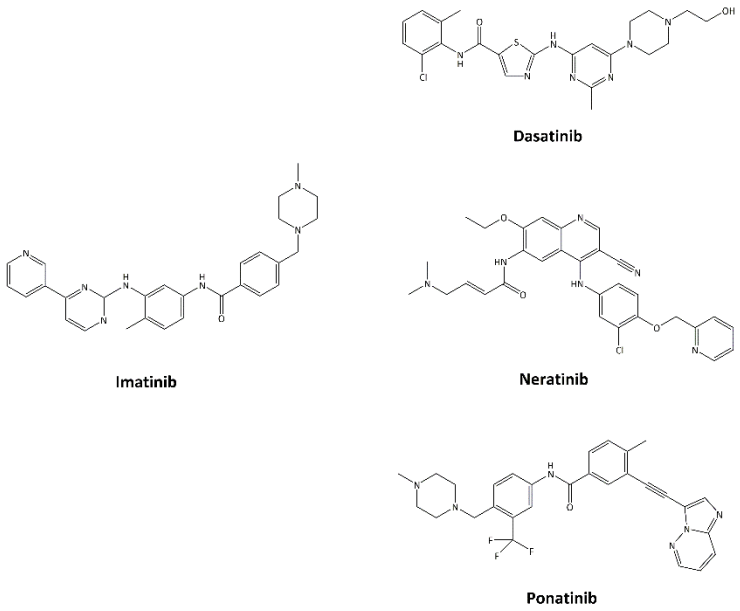


Figure 1.11. **Structures of first and second generation protein kinase inhibitors**. Imatinib was the first approved small-molecule inhibitor used to treat CML (Deininger *et al.*, 1997). Extensive drug resistance in patients in response to imatinib treatment led to the generation of next generation inhibitors to treat CML, including dasatinib, neratinib and ponatinib (Shah *et al.*, 2004; O'Hare *et al.*, 2009; Blay & von Mehren, 2011). All four inhibitors are designed to target the BCR-ABL fusion protein found in CML patients.

1.6.1. **Aurora & Polo-like kinases as drug targets**

Both Polo-like and Aurora kinases have been identified as drug targets, due to their key roles in regulating mitosis and frequent misregulation in cancer cells. A number of inhibitors designed against these kinases are currently undergoing clinical trials. PLK inhibitors in clinical development include Volasertib, Rigosertib (a dual PLK and PI3K inhibitor) and GSK461364A (Gilmartin *et al.*, 2009; Rudolph *et al.*, 2009; Casolaro *et al.*, 2013). Volasertib is an ATP-competitive inhibitor, which potently inhibits PLK1-3 and has reached phase III

clinical trials to treat leukaemias. Volasertib was shown to treat bladder cancers by inducing cell cycle arrest and apoptosis.

PLK4 was identified as a drug target following the results of an siRNA screen against protein kinases in breast cancers, which revealed reduced proliferation upon gene silencing of PLK4. A subsequent drug screen identified the ATP-competitive inhibitor, CFI-400945 as a potent inhibitor of PLK4 activity (Mason *et al.*, 2014). Inhibition with CFI-400945 prevented centriole over duplication and led to an accumulation of PLK4, resulting from an inability to trans-autophosphorylate within its degradation motif. PLK4 is the most structurally diverse of the polo-like kinase family and CFI-400945 does not inhibit PLK1-3. However, it does potently inhibit several other protein kinases, including Aurora B.

The critical role of Aurora kinases has resulted in the development of a number of inhibitors targeting the activity of these enzymes. These include the Aurora A inhibitors Alisertib and Danusertib currently in clinical trials, and Barasertib which inhibits Aurora B (Manfredi *et al.*, 2011; Qin *et al.*, 2015). In addition, VX-680 is a potent inhibitor of Aurora A, B and C and induces cell cycle arrest and apoptosis when used as a treatment for leukaemia, lymphoma and colorectal cancer (Harrington *et al.*, 2004) (Figure 1.12).

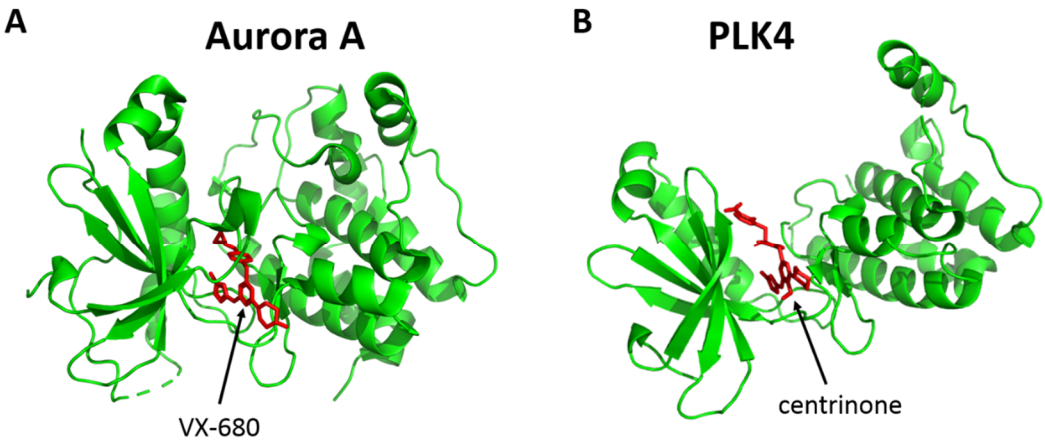


Figure 1.12. Protein kinases bound to small-molecule inhibitors. A) Aurora A inhibited by the ATP-competitive inhibitor VX-680. B) PLK4 inhibited by the ATP-competitive inhibitor, centrinone. Images produced in PyMOL.

VX-680 is also a potent inhibitor of PLK4 and was used as a template for the design of the novel PLK4 inhibitor, centrinone (Tyler *et al.*, 2007; Sloane *et al.*, 2010). Chemical alterations to VX-680 included the addition of a methoxy substituent at the C5 position, targeting Met91 in the PLK4 hinge loop region (Harrington *et al.*, 2004; Wong *et al.*, 2015) (Figure 1.12). Centrinone treatment of cancer cells resulted in the depletion of centrosomes and cell-cycle arrest in G1 via a p53-dependent mechanism, potentially by loss of the interaction between p53 and MDM2. Inhibition of cells containing a drug-resistant PLK4 (in which Gly95 important for selectivity was mutated to Leu) provided confirmation that the loss of centrosome phenotype was a direct result of PLK4 inhibition, as centrosome duplication proceeded in this cell line (Wong *et al.*, 2015). Centrinone is suggested to exhibit three orders of magnitude greater affinity for PLK4 over Aurora A, and is now frequently used in research laboratories to study the activities of PLK4 in cells.

1.7. Mass Spectrometry-based proteomics

Proteomics is defined as the study of the complete set of proteins expressed by a genome within a particular cell, at any given time (Wilkins *et al.*, 1996). Proteomics has become an invaluable tool for the assessment of the global role of protein kinases in co-ordinating distinct cellular processes. As protein kinases primarily function by signal transduction via the transfer of phosphate to substrates, the phosphorylation status (and other PTMs) of proteins within the cell must also be comprehensively investigated to gain a true insight into the activities of kinases within the cell. Mass spectrometry (MS)-based proteomics using liquid chromatography-tandem mass spectrometry (LC-MS/MS) has developed rapidly over the past 20 years and allows theoretically for an unbiased assessment of the global proteome/modified proteome under study.

1.7.1. **Tandem mass spectrometry**

MS for proteomics studies typically couples liquid chromatography (LC) separation with tandem MS. For the analysis of peptides, reversed phase high-pressure liquid chromatography (RP-HPLC) is most commonly used due to the high level of separation that can be achieved and the use of solvents compatible with downstream electrospray ionisation (ESI). In RP-HPLC, peptides are bound to a column packed with silica particles

containing alkyl chains, of which C18 is the most common. Bound peptides are eluted from the column based on their relative hydrophobicity with increasing concentrations of organic solvent.

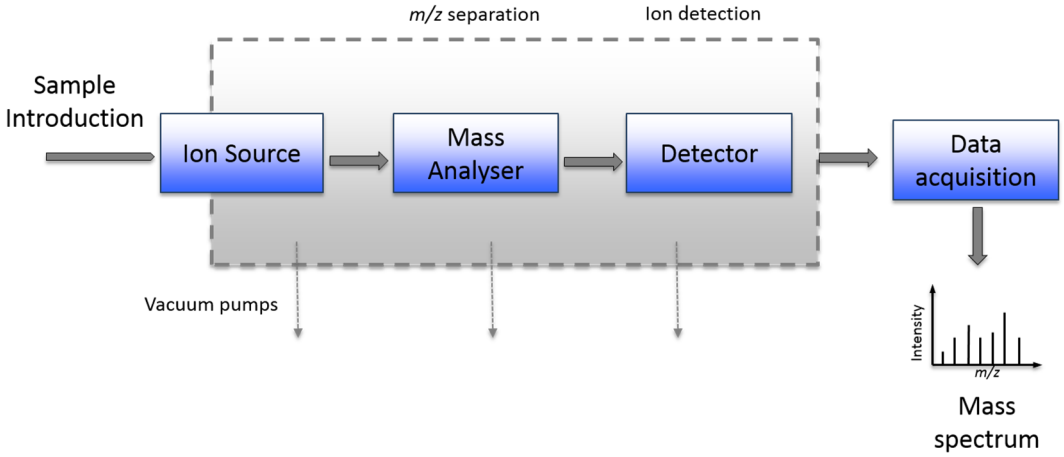


Figure 1.13. Schematic of a typical mass spectrometer. Ions are introduced in to the mass spectrometer following formation of gaseous ions at the source. Ions are resolved in the mass analyser based on their *m/z,* and detected. The information (*m/z*, relative abundance) is recorded by a data system. The grey box indicates the parts of the mass spectrometer that are operated under high vacuum.

As peptides elute from the reversed-phase column, they are first introduced to an ionisation source, which generates gaseous ions. The ions are then resolved based on their mass-to-charge ratio (*m/z*) and detected to produce a mass spectrum (Figure 1.13). In a typical tandem MS experiment required for proteomics studies, a survey scan (MS1), is first performed, to detect all ions of a defined *m/z* range at a given time and record their intensity. The mass of a peptide ion alone is insufficient for robust identification from a complex mixture, and therefore, peptide ions must be fragmented so that the constituent amino acids can be determined. In data-dependent acquisition (DDA) approaches, ions are selected for fragmentation based on their relative intensity, with the Top*N* most abundant ions being selected from the MS1 full scan (Mann *et al.*, 2001). These ions then undergo fragmentation and the product ions produced are detected to produce an MS2 spectrum. The quality of each spectrum determines the confidence in downstream identification by data processing software. One limitation of DDA experiments is the preclusion of low intensity precursors being selected for fragmentation and therefore a certain loss of information (Han *et al.*, 2008). An alternative, data-independent acquisition (DIA) approach, whereby all precursors over the defined *m/z* range are subjected to fragmentation, was developed to overcome this limitation (Chapman *et al.*, 2014). With both approaches,

peptide sequences are identified from the product ion spectra, and these peptide IDs are used to infer protein identity.

1.7.2. **Electrospray Ionisation**

Once ions have been introduced to the source, they must rapidly desolvate and enter the gas phase as charged ions in a process known as electrospray ionisation (ESI). ESI was first demonstrated by Fenn *et al*., in 1989 and later adapted into nanoelectrospray by Wilm & Mann in 1994 to use lower flow rates (Fenn *et al.*, 1989; S. Wilm & Mann, 1994).

In ESI, the sample is sprayed from a capillary subjected to a strong electric field and high temperature at atmospheric pressure. In positive-mode ESI, the sample is typically in an acidic solution and as it is sprayed from the capillary, the analytes move under the influence of the electric field (Kebarle & Verkerk, 2009; Wilm, 2011). The positively charged ions will move to the surface of the droplet and polarization causes the droplet to distort in to a cone shape, known as the Taylor cone (Taylor & McEwan, 1965). Instability at the cone results in droplets being emitted that undergo rapid desolvation until charged gas-phase ions are produced. Two models describe how charged gaseous ions are produced from a liquid sample: the ion evaporation model and charge residue model (Figure 1.14).

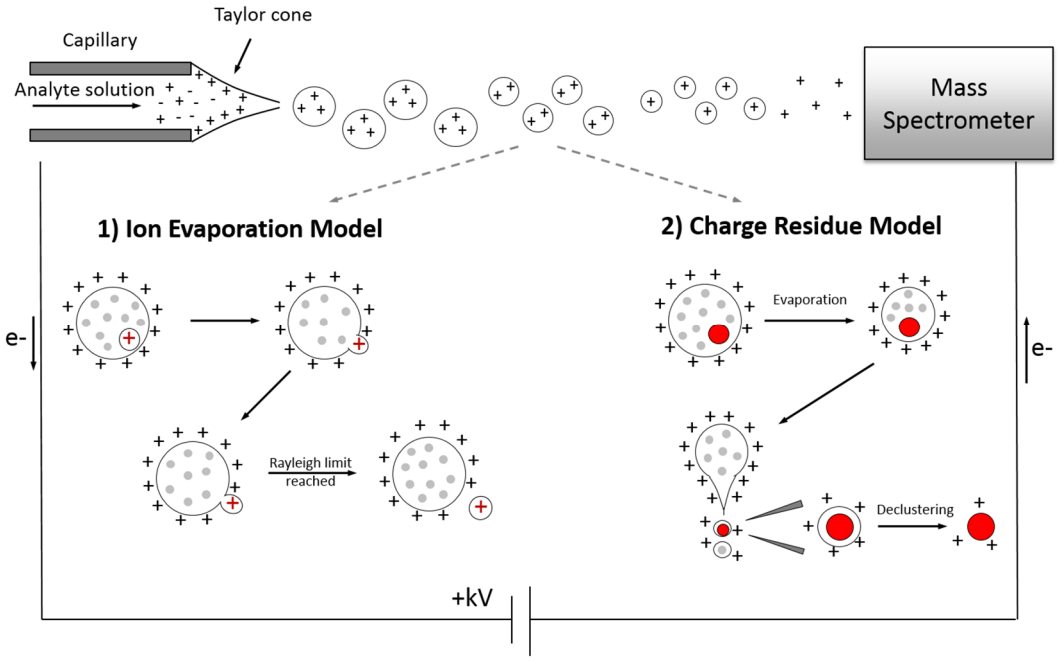


Figure 1.14. Electrospray ionisation. Two models describe the process of gaseous ion formation. 1) The ion evaporation model explains how smaller ions enter the gas phase (V. Iribarne & Thomson, 1976; Thomson & Iribarne, 1979) and 2) the charge residue model describes the process for larger molecules, including proteins (red circles) (Dole *et al.*, 1968). The grey circles depict peptides in the droplet. Figure adapted from (Wilm, 2011)

In the ion evaporation model, the emitted droplets undergo solvent evaporation, aided by heat. As the solvent evaporates and the water content of the droplet increases, the surface tension increases and charge density of the droplet builds up. Eventually, at what is defined as the Rayleigh limit, the repulsion between the charges exceeds the surface tension (V. Iribarne & Thomson, 1976; Thomson & Iribarne, 1979). At this point, Coulomb fission of the droplet occurs, leading to the ejection of solvated ions from the droplet surface. It is now widely accepted that this model explains ESI of small molecular weight molecules (Konermann *et al.*, 2013) (Figure 1.14).

The charge residue model is suggested to explain ESI of larger molecules, such as intact proteins (Dole *et al.*, 1968). This model assumes that ESI produces droplets containing one ion. As the droplet evaporates, a new Taylor cone is formed and emits smaller, highly charged droplets. Multiple iterations of this process continue until droplets containing a single ion are produced. These ions are released into the gas phase by desolvation and declustering (Figure 1.14).

1.7.3. **Thermo Orbitrap Fusion**

The Thermo Orbitrap Fusion is one of the most advanced mass spectrometers currently available, capable of achieving a resolution of 500,000 at 200 *m/z* (Figure 1.15). This 'tribrid' instrument consists of three different mass analysers: a quadrupole mass filter, ion trap and an orbitrap; a number of fragmentation modes including collision-induced dissociation (CID), higher-energy collisional dissociation (HCD) and electron transfer dissociation (ETD) are also permissible (Senko *et al*., 2013). Ions produced by ETD can be further activated by supplemental collision energy. This fragmentation strategy is described as ETcaD with low supplemental energy or EThcD if higher collision energies are applied. The mass analysers and fragmentation modes available on the Orbitrap Fusion will be discussed in detail below.

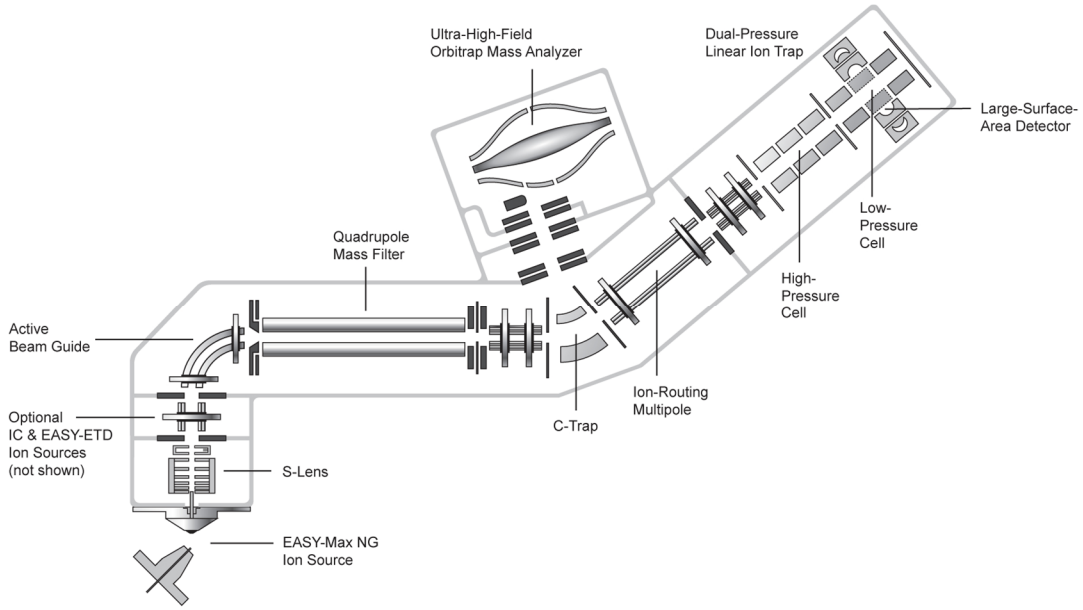


Figure 1.15. Schematic of the Orbitrap Fusion mass spectrometer. The Orbitrap Fusion Tribrid mass spectrometer includes a quadrupole mass filter, ion trap and Orbitrap mass analysers. CID and ETD can be performed in the linear ion trap, and HCD in the ion-routing multipole.

Following ESI in the Orbitrap Fusion, ions are passed through an S-lens, which increases sensitivity by focusing the ion beam and prevents unwanted mass separation. An active beam guide also helps to reduce noise by preventing neutrals from entering the quadrupole (Senko *et al*., 2013). For ion selection, an RF potential is applied to the quadrupole, enabling it to act as a mass filter, and transmit precursor ions of a defined *m/z* for analysis, with the *m/z* value transmitted being dependent on the voltages applied. Ions are then routed by

the ion-routing multiple which permits stable ion transfer between the multipole and both the linear ion trap and orbitrap mass analysers.

1.7.4. **Ion Trap**

Ion traps store ions using an oscillating electric field and can be used for mass analysis by measuring ion *m/z*. There are two types of ion trap: the 3D (or Paul) trap and the later developed 2D (or linear) ion trap which is implemented in the Orbitrap Fusion instrument (Schwartz *et al.*, 2002). The key advantages of linear ion traps over the Paul trap is reduced space-charge effects from an increased volume for ion storage, alongside higher trapping efficiencies which leads to greater sensitivity.

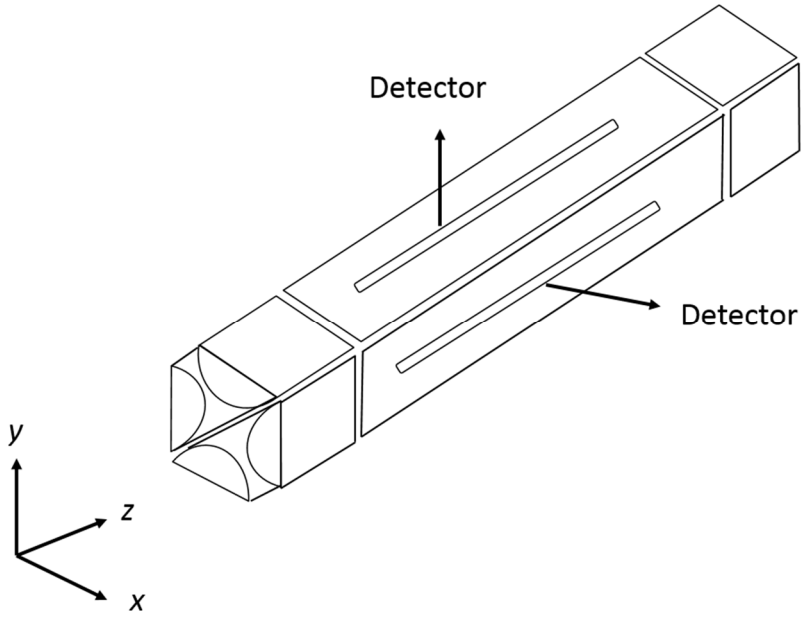


Figure 1.16. Linear ion trap mass analyser. The ions enter the trap and two sets of RF voltages cause the ions to oscillate in the radial direction, whilst a DC potential applied to the electrodes confines the ions axially. Ions exit the trap and are detected by one of two detectors. Adapted from Schwartz *et al*., 2002.

A linear ion trap is composed of four rods arranged around a central axis (Figure 1.16). When ions enter the ion trap, two sets of RF voltages are applied to the rods, with one set out of phase with the other. This confines the ions in the radial dimension, whilst ions are confined axially by a DC potential applied to the end electrodes. Ions are cooled by the injection of helium and mass analysis can be performed by the mass selective instability technique. This involves linearly ramping up the RF potential at one end of the electrode,

resulting in ion instability and ejection from the trap via two slots in opposite rods leading to two detectors for mass analysis (Stafford *et al.*, 1984; Douglas *et al.*, 2005).

Alternatively, mass analysis can be performed by resonant ejection, which enhances sensitivity and resolution. Resonant ejection exploits the relationship between resonance frequency and *m/z*. Specific resonance frequencies can be employed to select ions of specific *m/z* for ejection from the trap due to the increased amplitude of their oscillations (Goeringer *et al.*, 1992; March, 2012; Snyder *et al.*, 2017). This principle led to the ability to utilise ion traps for fragmentation and MS/MS analysis. By implementing frequencies that are resonant to all ions except the ion of interest, these ions will be ejected from the trap, leaving an ion of specific *m/z* that can be fragmented by collisional dissociation and mass analysed (Snyder *et al.*, 2017).

Although linear ion traps contain a larger volume for ion storage than 3D ion traps, the number of ions that enter the trap must still be controlled to prevent space-charging effects. As the number of ions increases, repulsion between ions increases (Vedel & André, 1984). This affects ion motion in the trap resulting in decreased resolving power and poor mass accuracy (Cox *et al.*, 1995).

1.7.5. **Orbitrap**

The Orbitrap mass analyser was developed by Alexander Makarov as a new approach for dynamic trapping of ions in an electrostatic field (Makarov, 2000). It consists of an outer 'barrel-like' electrode and an inner 'spindle-like' electrode (Figure 1.17). For ions to enter the orbitrap, they must first accumulate in the 'C-trap', an RF-only quadrupole, which is composed of four curved electrodes and located perpendicular to the orbitrap. The C-trap is filled with nitrogen bath gas to cool the ions prior to injection in to the orbitrap. The RF is ramped down and a high voltage is applied across the C-trap, which ejects ions of a specific *m/z* in to the Orbitrap (Makarov, 2000).

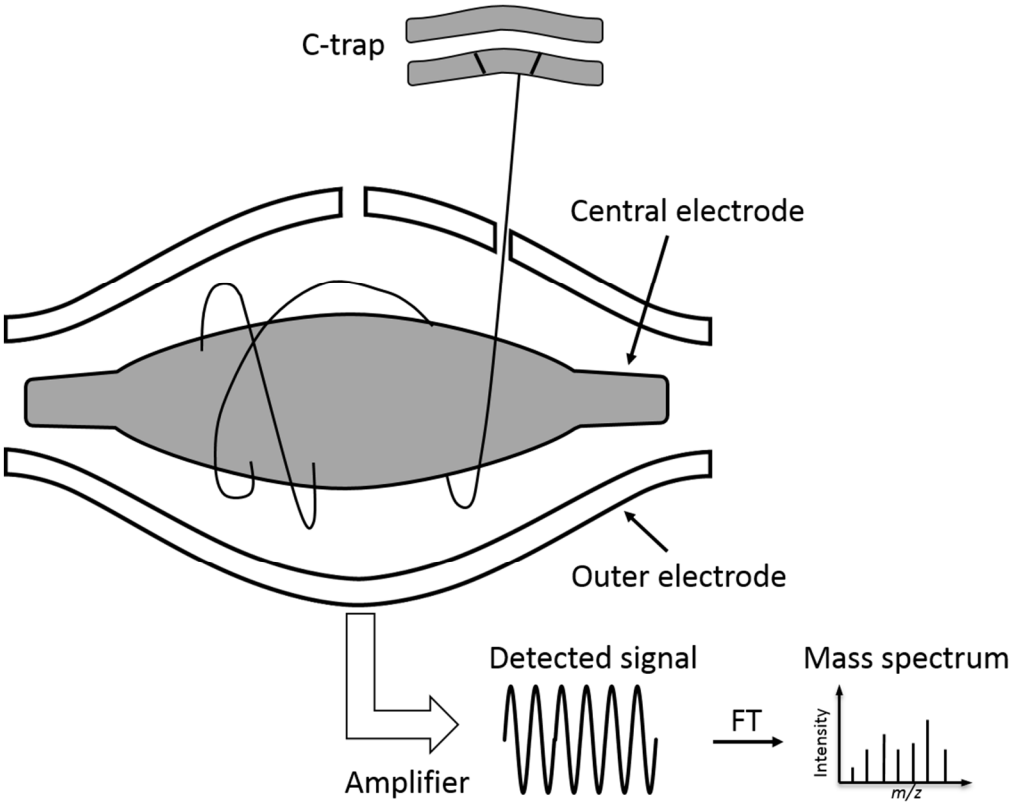


Figure 1.17. Orbitrap mass analyser. The ions are ejected from the C-trap and enter the orbitrap. Strong axial and radial electric fields result in the ions oscillating and orbiting around the central electrode. The frequency of oscillations is directly proportional to *m/z*. The ions are detected by the outer electrodes and the time-domain signal is Fourier Transformed (FT) to produce a mass spectrum. Figure adapted from Eliuk & Makarov, 2015.

Voltages applied between the two electrodes of the orbitrap result in strong axial and radial electric fields, which force the ions to oscillate in the *z*- direction, whilst simultaneously orbiting around the central electrode. To prevent ions from collisions with the outer electrodes, the voltages are increased, resulting in 'electrodynamic squeezing'. The electric field is strictly linear across the axis resulting in harmonic oscillations in this direction. The frequency of these oscillations is directly proportional to *m/z* and independent of the initial properties of the ion and is therefore the parameter used to determine mass. The outer electrodes serve as receiver plates to detect the current induced by the oscillating ions, as a 'time-domain' signal. This signal is then Fourier Transformed to produce a mass spectrum (Zubarev & Makarov, 2013; Eliuk & Makarov, 2015).

The implementation of the Orbitrap in commercial mass spectrometers has become a powerful tool in the proteomics field, as ions can be measured at high resolution (up to 1,000,000 at 200 *m/z*) and with high mass accuracy, significantly enhancing peptide identification confidence.

1.7.6. **Collision-induced dissociation (CID)**

In the Orbitrap Fusion, CID is performed in the high-pressure cell of the dual pressure linear ion trap (Figure 1.15). Precursor ions collide with the inert helium gas in the trap and the kinetic energy is transferred to vibrational energy. This vibrational energy re-distributes throughout the peptide ion, resulting in bond cleavage (Summerfield & Gaskell, 1997) (Figure 1.18). The 'mobile proton' theory suggests that peptide ion cleavage is charge-directed and that protons acquired during (n)ESI in positive ion mode localise at basic residues (Arg, Lys, His), but can be transferred upon activation to amide nitrogen atoms across the peptide. Consequently, the C-N amide bond is then cleaved, producing a series of *b*- and *y*- ions from the N- and C-termini respectively, depending on which terminus the charge is located (Roepstorff & Fohlman, 1984; Dongré *et al.*, 1996).

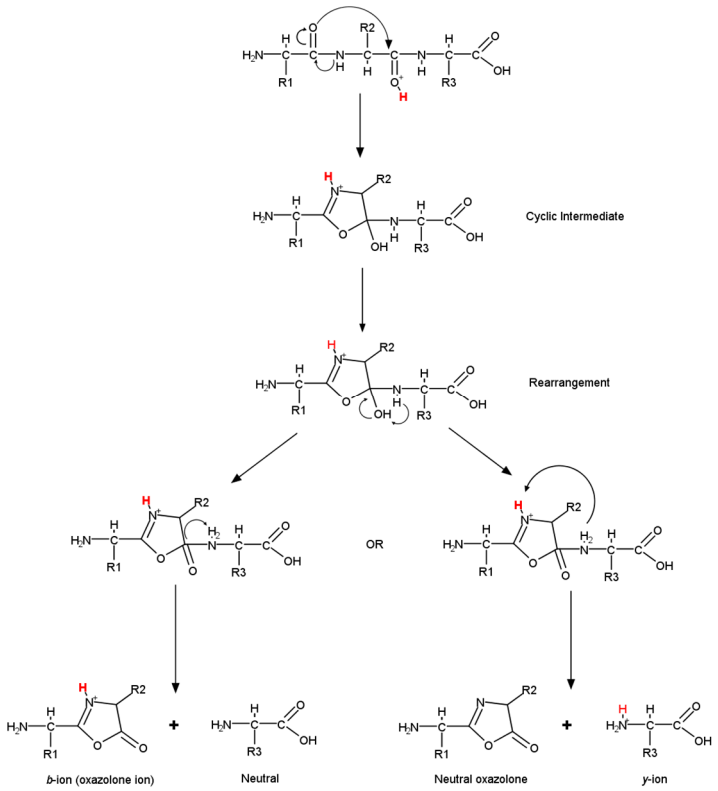


Figure 1.18. Collision induced dissociation (CID) mechanism of fragmentation. Peptide collisions with inert gas result in cleavage at the C-N bond, producing *b*- and *y*- ion series. Figure adapted from Wysocki *et al.*, 2000.

The efficiency of CID is strongly influenced by the residues present within the peptide. Basic residues (Arg, Lys, His) and their location within the peptide affect ionisation efficiency, charge state and consequently the fragmentation pathways and thus the CID product ion generated. Once a peptide has been cleaved, the fragment ions produced can form a temporary dimer to try to retain the proton therefore, basic residues with high proton affinity will influence fragmentation (Tabb *et al.*, 2004). In addition, proline exhibits high gas-phase basicity as it is a secondary amine with high proton affinity. Therefore, cleavage N-terminal to Pro is favourable and intense *y*-ions are typically observed in the resulting mass spectrum (Hunt *et al.*, 1986).

Once the peptide ion has been fragmented, no further collisions occur during resonant CID, as the product ions will no longer be resonant with the frequency being applied in the trap, limiting the number of product ions detected. In addition, CID in an ion trap is limited by the 'one-third rule' whereby fragment ions with an *m/z* less than one-third of the precursor will suffer from instability in the ion trap and be lost. Despite this, the speed and sensitivity of CID fragmentation in an ion trap has led to it being widely used for peptide analysis.

1.7.7. **Higher-energy collisional dissociation (HCD)**

HCD describes the fragmentation mode implemented within Orbitrap instruments and is broadly speaking, equivalent to the higher energy CID that occurs in a hybrid quadrupole/time-of-flight mass analyser. In HCD, ions are also fragmented upon collision with an inert gas. However, fragmentation occurs within a collision cell with high RF voltage, rather than in an ion trap. This has a number of advantages including no low mass cut-off, high resolution mass analysis and the potential for increased sequence coverage due to the application of higher energies than are used in CID. In addition, HCD in a collision cell allows the peptides to undergo multiple fragmentation events. HCD fragmentation can also produce additional fragments, such as immonium ions, which can aid peptide identification (Olsen *et al.*, 2007).

One limitation of HCD fragmentation is that acquisition times are much longer than that for CID in an ion trap due to a much greater number of ions being required for Orbitrap mass analysis (Jedrychowski *et al.*, 2011). This can impede the number of identifications. However, recent developments in Orbitrap platforms, such as the Orbitrap Q-Exactive HF-X

have much faster ion-transfer times enabling rapid MS/MS analysis and higher numbers of identifications in a shotgun proteomics experiment (Kelstrup *et al.*, 2018).

1.7.8. **Electron transfer dissociation (ETD)**

ETD fragmentation was developed in 2004 to overcome the limitations of electron capture dissociation (ECD) fragmentation, which requires FT-ICR instrumentation not commonly used in the proteomics field. Instead of using near thermal electrons as in ECD, ETD uses reagent anions for ion-ion reactions to transfer an electron to the peptide cation, resulting in peptide cleavage at the N-Cα bond and a series of 'even electron' *c*- and 'odd electron' *z*-type fragment ions (Syka *et al.*, 2004) (Figure 1.19).

The first step of ETD is the generation of reagent anions and can be achieved in a number of ways. Chemical ionisation (CI) sources produce near-thermal electrons from the collisions between electrons and an inert gas. Reagent anions (most commonly fluoranthene) can then capture the electrons, producing radical anions. Alternatively, a Townsend glow discharge can be used to produce thermal electrons used to generate radical anions and is utilised in both the Orbitrap Fusion and Lumos instruments. The reagent source is located at the front end of the instrument, between the S-Lens and Active Beam Guide. Switching polarity of the instrument allows the radical anions to be *m/z* selected and transferred to the high pressure cell of the ion trap for the ETD reaction.

Transfer of an electron to the peptide is a non-ergodic process and thus, the energy does not distribute throughout the ion in the same way as during collisional dissociation. Therefore, labile modifications (such as phosphorylation) can be maintained on the modified amino acid which greatly aids phosphosite localisation. Electron transfer results in charge neutralisation and therefore, ETD requires multiply protonated peptides for cleavage. Peptides with low charge form compact structures in the gas-phase whereas ions with greater charge density exhibit linear gas-phase structures (Riley & Coon, 2018). The application of ETD to peptides with compact structures often leads to non-covalent interactions between the resulting *c*- and *z*- ions, greatly reducing sequence information in the mass spectrum (Good *et al.*, 2007; Liu & McLuckey, 2012).

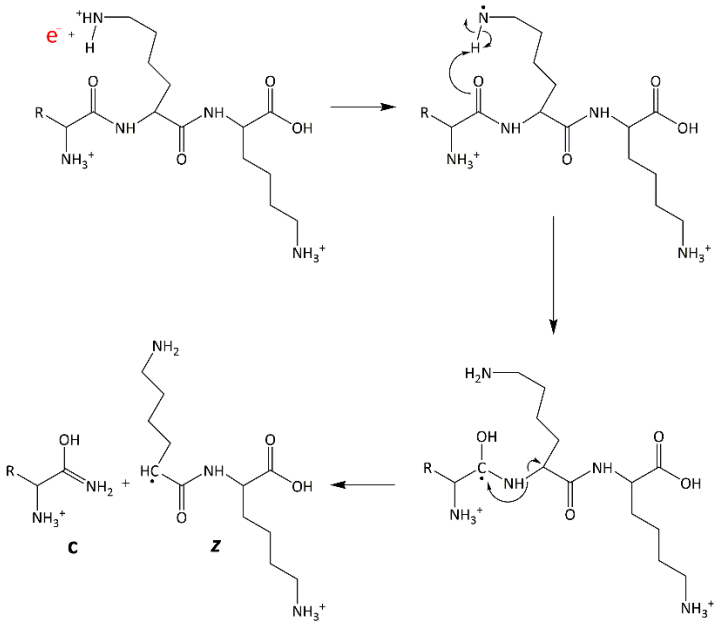


Figure 1.19. Electron transfer dissociation (ETD) mechanism of fragmentation. Electrons are donated by radical fluoranthene anions. This results in cleavage at the N-Cα bond, producing *c*- and *z*- ion series. Adapted from Syka *et al.*, 2004.

Strategies to overcome non-covalent interactions between fragment ions include the use of supplemental energy to dissociate these interactions and generate richer product ion spectra to aid peptide identification. Following ETD, the unreacted and charge reduced precursor and non-dissociative product ions can be transferred to the collision cell for additional ion activation. If low supplemental energy is applied (NCE <15), then the resulting mass spectrum will contain a richer *c*- and *z*- ion series than that of ETD alone (Swaney *et al.*, 2007). However, if higher collision energies are applied, then in addition to dissociating ETD product ions, additional fragmentation via collisions with the inert gas will occur, leading to spectra containing *c*-, *z*-, *b*- and *y*- ion series. This generates much greater sequence coverage, aiding peptide identifications (Frese *et al.*, 2012; Frese *et al.*, 2013)

ETD efficiency is also influenced by peptide sequence. In contrast to CID, ETD is unable to cleave peptides N-terminal to proline residues. This is because the cyclic secondary amine structure of the proline side chain would require two N-Cα bonds to be broken to produce the fragment ion (Kim & Pandey, 2012). In this way, CID and ETD can be used as complementary techniques to improve sequence coverage.

1.7.9. **Thermo Orbitrap Fusion flexibility**

The advanced capabilities of the Orbitrap Fusion allow for deep coverage of the proteome and phosphoproteome, which will be required to study the global activities of protein kinases discussed in this thesis. The ion trap and orbitrap mass analysers can operate in parallel such that a full-scan can be performed in the orbitrap, whilst ions from the previous full scan are fragmented and detected in the ion trap. Previous Orbitrap instruments have employed a Top*N* approach to select ions for fragmentation. However, the Orbitrap Fusion utilises a 'Top Speed' approach where ions are fragmented over a pre-determined cycle time. This approach allows for potentially larger numbers of identifications without extending the overall duty cycle.

Flexibility at all stages, including the selection of mass analyser to be used for ion detection and fragmentation mode to be utilised, ensures that the setup of system can be customised for specific applications.

1.7.10. **Proteomics workflows**

The most widely utilised strategy for proteomics analysis is a 'bottom up' approach, in which proteins are first digested in to peptides for MS analysis. This strategy is summarised in Figure 1.20. First, samples are lysed using an appropriate lysis buffer to release the proteins within the cells. The disulphide bonds maintaining protein tertiary structure are reduced using dithiothreitol (DTT) or tris(2-carboxyethyl)phosphine (TCEP). To prevent disulphide bonds from re-forming, samples are incubated with iodoacetamide or iodoacetic acid to alkylate free cysteine sulfhydryl groups. Samples must then be digested to form peptides of a suitable size for analysis, with trypsin the most common choice. Trypsin cleaves C-terminal to Lys and Arg residues (except when the residue is followed by a Pro residue) and produces peptides of suitable length for C18 reversed-phase separation and MS analysis (7-20 amino acids). In addition, the basicity of Lys and Arg means that all tryptic peptide ions will carry an extra proton, in addition to the proton localised to the NH_2- group of the N-terminus following ESI.

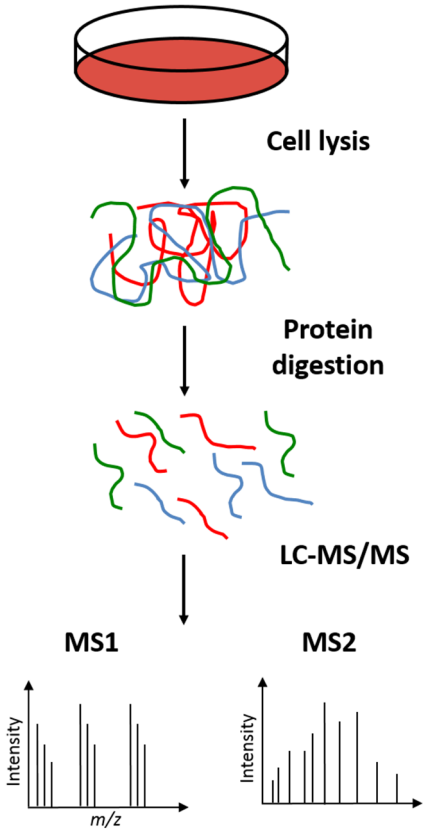


Figure 1.20. Bottom-up proteomics workflow. Proteins from a biological sample are digested to form peptides and analysed by LC-MS/MS

1.8. Quantification strategies

Typically, large-scale proteomics analyses aim to not only identify which proteins are expressed within a cell, but also to quantify the relative or absolute abundances of the identified proteins. Often, differences in protein expression arising from specific cell stimuli, or in response to drug treatments, are required to understand biological processes.

1.8.1. **Absolute quantification**

Absolute quantification reports exactly how much of a given protein is expressed, and this is typically reported as number of copies per cell, or moles per cell. Intensity-based absolute quantification (iBAQ) is commonly employed for measuring how much protein is in the sample. In iBAQ, the total intensities of the observed peptides are summed and divided by the number of theoretically observable peptides of a protein (Schwanhausser *et al.*, 2011). The iBAQ scores can then be converted in to number of copies of a particular protein

per cell. Alternatively, absolute quantification can be performed by adding known amounts of an isotopically labelled identical analyte standard, such as chemically synthesised AQUA peptides or peptide concatamer QconCAT standards. Known amounts of standard are spiked in to samples prior to MS analysis, allowing the amount of peptide (and therefore protein) within a sample to be determined (Beynon *et al.*, 2005; Rivers *et al.*, 2007; Kettenbach *et al.*, 2011a).

1.8.2. **Relative quantification**

Relative quantification of peptides can be performed using label-based or label-free approaches. In label free quantification, each sample to be measured (e.g. control vs treated) is analysed in a separate experiment (Figure 1.21). If carefully controlled, the retention times for ions will be the same across runs and can be aligned. The signal intensity (or peak area) is then measured for each peptide and the areas for peptides of a particular protein are averaged. The areas are then compared to proteins in different samples to assess relative changes (Bondarenko *et al.*, 2002). Spectral counting can also be used to quantify proteins in a label-free experiment. In this approach, the number of MS2 spectra acquired for a given peptide is summed to determine relative protein abundance. This is based on the observation that the number of MS2 spectra acquired and the peptides identified for a protein increases as protein abundance increases (Liu *et al.*, 2004a) This can then be compared to different samples to assess relative changes in protein abundance. This is because in DDA experiments, an abundant peptide will be sampled more frequently. This approach has a number of inherent limitations including the stochastic nature of DDA reducing reproducibility and the inherent bias toward abundant proteins (Lundgren *et al.*, 2010).

Labelling strategies for relative quantification include metabolic approaches to label proteins as they are synthesised within the cell (*i.e.* SILAC), or chemical approaches to label digested peptides.

Stable-isotope labelling by amino acids in cell culture (SILAC) is the most popular labelling method and widely utilised in quantitative proteomic workflows. SILAC involves supplementing culture medium with stable isotopes of amino acids, which are incorporated in to the proteome through normal metabolic processes (Ong *et al.*, 2002; Ong & Mann,

2007). Heavy lysine ($^{13}C_6$, $^{15}N_2$) and arginine ($^{13}C_6$, $^{15}N_4$) are most commonly used, as peptides for LC-MS/MS analysis are produced following tryptic digestion at these residues. When proteins from heavy and light samples are mixed 1:1 and analysed by LC-MS, the peptides appear as a pair in the mass spectrum, separated by the difference in mass corresponding to the incorporated stable isotope (Figure 1.21). The ratio of signal intensities can then be compared to provide quantitative information on the relative protein abundance in the cell. SILAC-based quantification ideally requires cells to be fully labelled prior to sample mixing, as partially labelled cells will contribute to the 'light' peptide signal upon analysis of the resulting peptides by MS, resulting in inaccurate quantification. SILAC labelling was the strategy used in this thesis for peptide and protein quantification.

As an extension of SILAC, a super-SILAC approach was developed in which a SILAC sample is prepared separately and spiked in to each of the samples to be analysed. The ratio between the 'light' peptides and the SILAC standard can be determined and as the amount of spiked in standard is equal across samples, a 'ratio of ratios' can be performed (Geiger *et al.*, 2011).

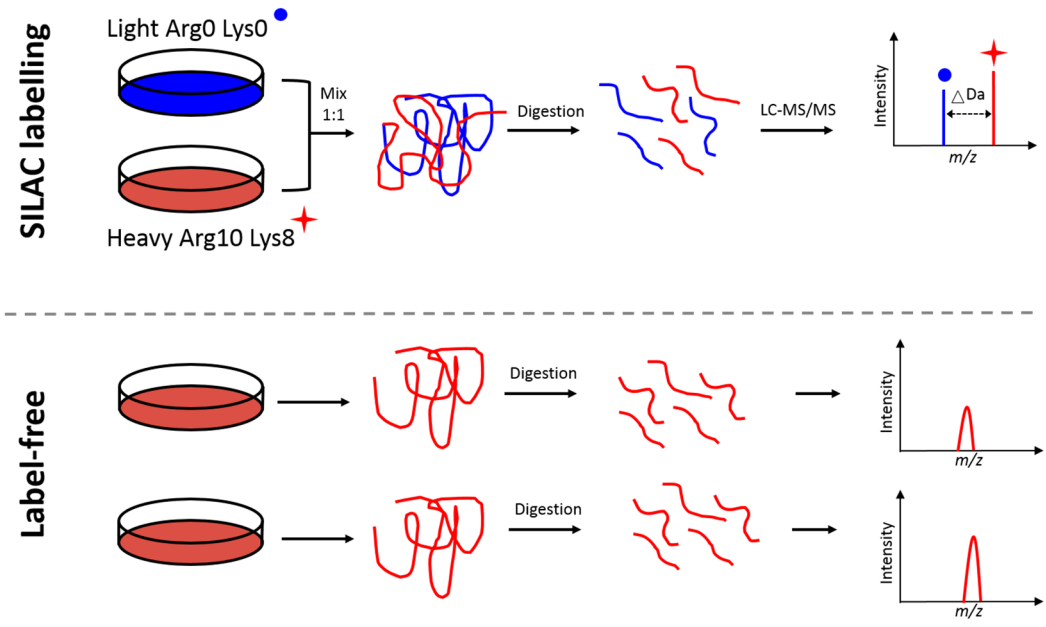


Figure 1.21. Comparison of label-based and label-free quantitative proteomics workflows. The upper panel shows a typical SILAC workflow. Cells are cultured in medium containing heavy or light stable amino acid isotopes. Samples are mixed 1:1, digested and analysed by LC-MS/MS. SILAC pairs are identified by a predefined difference in mass and quantified based on MS1 intensity. The lower panel shows a label-free workflow where samples are prepared separately and analysed in parallel. The peaks are quantified by comparing extracted ion chromatograms.

Chemical labelling approaches include stable-isotope dimethyl labelling, or isobaric labelling with iTRAQ or TMT reagents. Dimethyl labelling uses formaldehyde to convert all primary amines (N-terminus and Lys side chains) to dimethyl amines. The resulting tryptic peptides will therefore have one dimethyl group for Arg terminating peptides and two for Lys terminating peptides, generating mass shifts of 4 Da and 8 Da respectively (Hsu *et al.*, 2003; Boersema *et al.*, 2009b). These mass shifts are then used to determine the ratio between 'heavy' and 'light' samples. Labelling with iTRAQ or TMT reagents use isobaric amine reactive tags to label peptides. Up to 8 iTRAQ labels and 11 TMT labels can be used in multiplex experiments with no difference in mass therefore; quantification is performed at the MS2 level. Upon peptide fragmentation, a reporter ion is released from the tags, each with a different mass. The relative intensities of each can be used to determine relative abundance of the peptide in each sample analysed (Thompson *et al.*, 2003; Ross *et al.*, 2004)

1.9. Phosphoproteomics

Phosphoproteomics is the study of the phosphorylation status of all of the proteins expressed within a cell. The increase in mass (80 Da) of a phosphorylated peptide allows modified peptides to be distinguished from their non-phosphorylated counterparts. Fragmentation of phosphopeptides can isolate the exact site of modification if sufficient coverage of the peptide sequence is obtained. However, analysing the phosphoproteome is a considerable challenge, with many 'bottlenecks' that need to be overcome.

One key issue is that phosphopeptide ions subjected to collisional dissociation can lose the attached phosphate group during fragmentation, a process termed 'neutral loss'. Losses of 80 Da, corresponding to HPO_3 or 98 Da, corresponding to H_3PO_4 can occur, with loss of 98 Da most commonly observed for pSer and pThr containing peptides. Neutral loss is most frequently observed when phosphopeptides are fragmented with CID and can be explained by the mobile proton model. In a mobile proton environment (fewer basic residues than the number of acquired protons), the protons can transfer to the negatively charged phosphate group, making the β-carbon of the side chain more electrophilic, which can then undergo intramolecular nucleophilic substitution leading to neutral loss of H_3PO_4 (Figure 1.22 A). In a limited mobile proton environment, neutral loss can also occur, due to the lower energy of the phosphoester bond compared with the amide bond. In this

environment, protons are localised to basic residues in the peptide, which can interact with the phosphate group and drive preferential cleavage of the phosphoester bond (Figure 1.22 B) (Palumbo *et al.*, 2008; Boersema *et al.*, 2009a; Lanucara *et al.*, 2014). The aromatic ring of pTyr prevents neutral loss of H_3PO_4 and loss of HPO_3 is primarily observed (Tholey *et al.*, 1999). CID in an ion trap often results in a mass spectrum dominated by the precursor exhibiting neutral loss, as no further fragmentation of the peptide will occur as the CID product ions are no longer resonant.

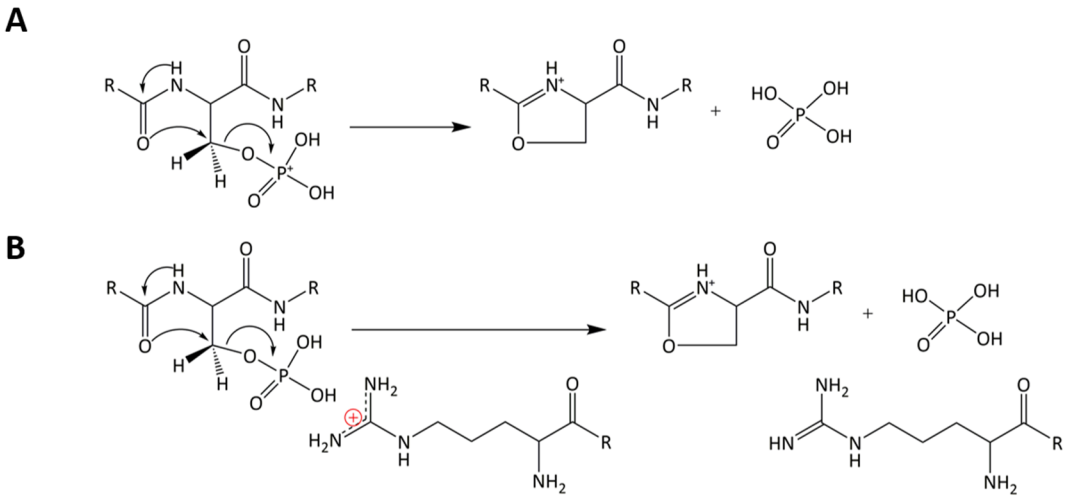


Figure 1. 22. Proposed pathways for neutral loss of H_3PO_4 from a phosphorylated peptide. A) Neutral loss in a mobile proton environment via an intramolecular nucleophilic substitution reaction. B) Neutral loss in a limited mobile proton environment via a charge directed nucleophilic substitution reaction. Adapted from Lanucara *et al.*, 2014.

Methods to improve fragment ion spectra acquired through CID fragmentation have included the development of multistage activation (MSA) methods. In MSA, the phosphorylated precursor is first activated, which will produce a predominant neutral loss fragment. This fragment then undergoes further dissociation in the ion trap and a pseudo MS^3 spectrum is produced, which combines the MS2 and MS^3 fragmentation spectra. This leads to richer fragmentation spectra and maintains fast acquisition speeds (Schroeder *et al.*, 2004).

Neutral loss can also be observed with HCD fragmentation. However, the ability for multiple collisions leading to additional fragmentation events can produce rich fragmentation spectra, allowing for the modified site to be identified. The improvements in

instrumentation have led HCD fragmentation and high resolution mass analysis to begin to overtake CID (and MSA) methods as the method of choice for phosphoproteomics analysis. Alternatively, ETD fragmentation can be utilised for phosphopeptide analysis, as the nonergodic nature means that the phosphate group is retained on the modified residue, often allowing the site of modification to be identified with greater confidence. However, ETD is only available on selected MS platforms, requires longer reaction times, and fragmentation is generally much less efficient than collision-mediated dissociation, in particular for low charge states (where $z = 2$). This has limited the benefit of ETD in high-throughput phosphoproteomics studies.

An additional challenge to performing phosphoproteomics is that phosphorylation is sub-stoichiometric and phosphoproteins are present in a background of highly abundant non-modified proteins (Steen *et al.*, 2006). DDA approaches to study phosphorylation will select the Top*N* most abundant ions, which will greatly limit the number of phosphopeptides selected for fragmentation. In addition, phosphopeptides are believed to exhibit reduced ionization efficiencies and can suffer from ion suppression (Ishihama *et al.*, 2007; Marcantonio *et al.*, 2008). Improvements in phosphopeptide analysis have therefore involved both offline fractionation methods (*e.g*. strong anion exchange, HILIC) to enable analysis of greater numbers of low abundant ions, and enrichment of phosphopeptides to separate them from non-phosphorylated peptides for analysis.

1.9.1. **High pH reversed phase fractionation**

Offline peptide fractionation has greatly aided the ability to identify large numbers of phosphorylated peptides in a sample. A number of methods have been utilised, including ion exchange (IEX) (Beausoleil *et al.*, 2004), hydrophilic interaction chromatography (HILIC) (McNulty & Annan, 2008), electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) (Alpert, 2008) and high pH reversed phase (high pH RP) (Yang *et al.*, 2012; Batth *et al.*, 2014). High pH RP was utilised in the work described in this thesis and is therefore described in further detail.

High pH RP fractionation separates peptides based on their relative hydrophobicity, with peptides of increasing hydrophobicity eluting as the concentration of solvent increases. The high pH of the buffers enables peptide separation that is orthogonal to the low pH RP

employed for separation of peptides prior to MS/MS analysis. At high pH, the charge distribution within each peptide is radically changed, as the carboxylic and ammonium groups within the peptide will be deprotonated (Gilar *et al.*, 2005; Delmotte *et al.*, 2007)). This will affect the elution profile compared to low pH RP and enhances the separation of peptides to increase overall phosphopeptide identifications. Compared with other fractionation strategies, high pH RP has shown to significantly improve phosphopeptide identifications, with more than double the number of phosphopeptides identified compared with strong cation exchange (Batth *et al.*, 2014).

1.9.2. **Titanium dioxide enrichment**

In addition to off-line fractionation, enrichment of phosphopeptides is a prerequisite to perform in-depth phosphoproteomics analysis. Traditionally, immobilised metal affinity chromatography (IMAC) has been widely used for phosphopeptide enrichment. In IMAC, the affinity of immobilised metal ions (e.g Fe^{3+}) for negatively charged groups is exploited (Neville *et al.*, 1997). Whilst successfully employed in many experiments, IMAC is limited by the affinity for acidic carboxylic acid groups on non-phosphorylated peptides to bind to the metal ions (Negroni *et al.*, 2012). These limitations led to the development of novel approaches for enrichment, including metal oxide affinity chromatography (MOAC), which most commonly utilises TiO_2 to bind to phosphopeptides.

TiO_2 enrichment is performed at low pH, which helps prevent the unwanted binding of acidic groups from non-phosphorylated peptides (Larsen *et al.*, 2005). Additives such as glutamic acid can also also be included, to further prevent carboxylic acid groups of non-phosphorylated peptides from binding (Wu *et al.*, 2007). The oxygen group of the phosphate has a strong affinity to the titanium metal, which binds via a bridging bidentate interaction (Figure 1.23). Phosphopeptides can then be eluted with an alkaline buffer such as ammonium hydroxide, which disrupts this interaction (Thingholm *et al.*, 2006).

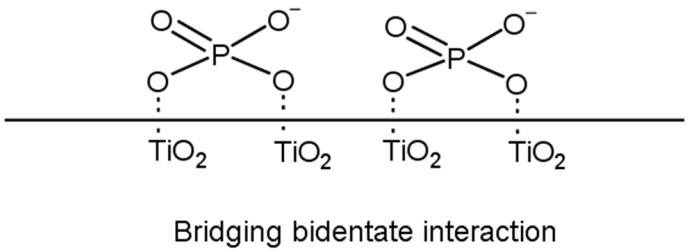


Figure 1. 23. Titanium dioxide enrichment of phosphopeptides. Phosphorylated peptides bind to the titanium dioxide resin via the formation of a bridging bidentate interaction between the metal and phosphate. Adapted from Larsen *et al.*, 2005.

Combining off-line fractionation with selective phosphopeptide enrichment has greatly advanced the study of complex phosphoproteomes by LC-MS/MS. Phosphoproteomics experiments can now identify tens of thousands of phosphopeptides and deliver novel insights in to the dynamic regulation of phosphorylation within cells. This includes studies in to the deregulation of phosphorylation in disease, as well phosphorylation dynamics in response to drug treatments. One study, which aimed to comprehensively analyse the phosphoproteome of two cancer cell lines (HeLa and K562) identified 16,000 and 24,000 phosphosites respectively (Zhou *et al.*, 2013).

1.10. Data Analysis

Following LC-MS/MS analysis, the acquired data must be computationally deconvoluted and processed to determine peptide identity based on the MS1 and MS2 spectra. Many search engines are available, including MASCOT, SEQUEST, Andromeda, PEAKS and MS Amanda (Eng *et al.*, 1994; Perkins *et al.*, 1999; Cox *et al.*, 2011; Dorfer *et al.*, 2014). The search engines typically use the experimental data (MS1 and MS2) and compare the masses to calculated theoretical values based on *in silico* digests of a proteome. A probability based scoring algorithm is used to determine confidence in the matched peptides (Perkins *et al.*, 1999). There is a requirement to determine the false discovery rate (FDR) of peptides identified to filter out false positive results. There are different approaches to calculating FDR, with a 'target decoy' approach being the most common. This approach reverses the sequences present in the database used to search data and the number of matches to the reversed database is an indicator of the number of false positives in the data (Elias & Gygi, 2007).

Additional software is required if peptide/protein quantification is required. The two most commonly used platforms are Progenesis for label-free data and MaxQuant, which can perform both label-based and label free quantification, and was used in the work described in this thesis.

MaxQuant is freely available software and is the most widely used platform for processing SILAC data (Cox & Mann, 2008). MaxQuant uses feature detection (*m/z* and intensity of the peptide peaks) in each MS scan for quantification and fits Gaussian peak shapes using the number of data points per peak. Ion retention time, *m/z* and intensity are then used to create a 3D peak, which is smoothed and used to determine an intensity weighted mass. In SILAC experiments, MaxQuant detects heavy and light SILAC pairs if they have the same charge and differ by a predefined mass (set in the method as a variable modification) (Cox & Mann, 2008). Peptides are identified using the Andromeda search engine implemented in to MaxQuant and FDR calculated using a target-decoy approach (Cox *et al.*, 2011). The advanced algorithm for quantification and ease of use make MaxQuant a popular choice for data analysis. In addition, output files from MaxQuant contain all of the information required to interrogate quantitative changes between samples to study biological processes (Cox & Mann, 2008).

1.10.1. **Phosphosite localisation**

The ability to identify large numbers of phosphopeptides has greatly advanced the phosphoproteomics field. However, to truly understand phosphorylation dynamics, the exact site of modification within a peptide must be determined. Many bioinformatics tools have been developed to apply calculated localisation scores to phosphosites within a given peptide sequence. These include A-score, PTM-score, SloMo, LuciPHOr and ptmRS (Beausoleil *et al.*, 2006; Olsen *et al.*, 2006; Bailey *et al.*, 2009; Taus *et al.*, 2011; Fermin *et al.*, 2013). Each of these tools aims to localise phosphosites by calculating the chance of a given peak (capable of being phosphorylated) being matched at random. An alternative approach is to compare search engine scores calculated when the phosphosite is localised at different residues within the peptide. This is incorporated in the MASCOT search engine and is referred to as a MASCOT delta score (Savitski *et al.*, 2011).

PTM-score determines sites of modification by first dividing the spectrum in to separate *m/z* 'bins'. The top *N* most intense ions within each bin are then used to score the probability of a residue being modified. One limitation of this approach is that ions of low intensity in bins containing few ions will be used in the calculation, which can affect the ability of the algorithm to determine true sites of modification. PTM-score calculates a probability score for the peptide with each of the phosphorylatable residues chosen as modified site. The scores are then converted to ensure the probabilities equal 100 % (Olsen *et al.*, 2006). PTM-score was developed for low-resolution CID data and therefore, the strength of the algorithm to correctly localise phosphosites in high-resolution data is likely limited. The most recently developed phosphosite localisation tool, ptmRS, is the latest development of the PhosphoRS algorithm and is integrated in the Proteome Discoverer platform (Taus *et al.*, 2011). It can be therefore be used with data searched in MASCOT and SEQUEST. The scoring in *ptmRS* was developed for high and low resolution data, acquired using a range of fragmentation modes (CID, HCD, ETD, EThcD). Compared with PTM-score, *ptmRS* considers the total number of extracted peaks across the full mass range of the MS2 spectrum, overcoming potential issues of uneven peak distribution in individual *m/z* bins (Chalkley & Clauser, 2012).

The ability to identify high numbers of phosphopeptide identifications using the search engines described, along with the ability to assign phosphosite localisation scores to residues within a peptide has advanced the phosphoproteomics field. Large-scale phosphoproteomics studies can now routinely be performed to study a diverse range of biological systems with deep coverage of the post-translationally modified proteome.

1.11. Aims and objectives

The aim of this work is to study the PLK4 regulated phosphoproteome to identify cellular processes and signalling pathways regulated by PLK4 activity and potentially identify novel substrates. The full workflow, including development of biochemical tools to study PLK4 in human cells, and the optimisation of MS parameters for phosphopeptide analysis will be established. These tools will then be utilised to perform a large-scale SILAC-based quantitative phosphoproteomics experiment. Finally, this work aims to establish a novel interaction between PLK4 and Aurora A, to aid understanding of the roles of these kinases in co-ordinating activities during the cell cycle.

The aims will be met through the following objectives:

- Production and *in vitro* characterisation of recombinant WT and G95R-PLK4 proteins
- Generation of stable isogenic U2OS cells lines transfected with FLAG-WT or G95R- PLK4 full-length plasmids
- Assessment of optimal conditions for PLK4 inhibition with centrinone
- Optimisation of MS acquisition methods on the Orbitrap Fusion for large-scale phosphopeptide analysis
- Performing a large-scale SILAC based quantitative phosphoproteomics experiment to investigate PLK4 regulated signalling pathways
- Assess a possible interaction between PLK4 and Aurora A

2. Methods:

2.1. Reagents

Unless otherwise stated, general lab reagents were purchased from Sigma Aldrich.

2.2. Antibodies

The following antibodies were used: anti-PLK4 clone 6H5 (Millipore, 1/100 for western blot; immunofluorescence), anti-Aurora A (Cell Signalling Technologies, 1/5000 for western blot, 1/500 for immunofluorescence), anti-pericentrin (Abcam, 1/500 for immunofluorescence), anti- γ-tubulin (1/500 for immunofluorescence), anti-TAT1 (1/5000, for western blot), anti-GAPDH (1/5000, for western blot), anti-p21 (Abcam, 1/300 for western blot), anti-CCND2 (1/100 for western blot), anti-EGFR (1/500 for western blot), anti-pThr669 EGFR (1/100 for western blot).

2.3. DNA sample preparation

2.3.1. Polymerase Chain Reaction (PCR) amplification

Full length (1-2913) PLK4 insert (50 ng) was incubated with 1.5 mM $MgSO_4$, 200 µM dNTPs, 2.5% DMSO, 1 µM forward primer incorporating a Kozak sequence, FLAG tag and *Hind III* and *Not I* restriction sites at the 5' and 3' ends respectively, 1 µM reverse primer, 1x KOD buffer and 1 µL KOD polymerase. The PCR cycle was set up as follows: 95 °C for 2 minutes followed by 30 cycles of 95 °C DNA denaturation for 20 seconds, 64 °C primer annealing for 10 seconds and 70 °C extension for 2 minutes, with a 10 minute 70 °C hold after the last cycle.

2.3.2. **Agarose gel electrophoresis**

PCR product was analysed by agarose gel electrophoresis using a 1 % (w/w) agarose gel containing 5 µL Midori green nucleic acid stain and run at 100 V for 30 minutes. DNA bands were visualised by UV light.

2.3.3. **DNA gel extraction**

DNA bands were excised for DNA purification using the QIAGEN QIAquick Gel Extraction Kit following the manufacturers protocol. Gel bands were dissolved in buffer QG (volume 3x that of the gel band) at 50 °C for 10 minutes followed by the addition of isopropanol (volume 1x that of the gel band). DNA was collected in a QIAquick spin column at 15,000 *x g* for 1 minute. The column was washed with 500 µL buffer QG followed by two washed with 750 µL buffer PE and centrifuged at 15,000 *x g* after each step. DNA was eluted from the column by the addition of nuclease free H_2O and DNA concentration determined by nanodrop.

2.3.4. **Restriction digestion & ligation**

Not I and *Hind III* restriction enzymes were incubated with FLAG-PLK4 PCR product and separately, with pcDNA5 frt/to vector for 2 hours at 37 °C Enzymes were heat deactivated at 65 °C for 20 minutes and pcDNA5 was treated with alkaline phosphatase to prevent re-ligation of the plasmid.

Ligation reactions were set up with 3.5 fold excess PLK4 insert to pcDNA5 vector and incubated with 20,000 units/mL T4 ligase and 1x reaction buffer for 15 minutes at room temperature. T4 ligase was heat deactivated at 65 °C for 20 minutes.

2.3.5. **Site Directed Mutagenesis**

Full length wild type (WT) PLK4 pcDNA5 plasmid (6ng) was incubated with 1.5 mM $MgSO_4$, 200 µM dNTPs, 2.5% DMSO, 1 µM forward primer incorporating the mutation GGA (Glycine) to CGA (Arginine), µM reverse primer, 1x KOD buffer and 1 µl KOD polymerase. The PCR cycle was set up as follows: 95 °C for 2 minutes followed by 35 cycles of of 95 °C

DNA denaturation for 20 seconds, a temperature gradient between 55 °C and 65 °C for primer annealing and a 70 °C extension for 16 minutes.

2.3.6. **Bacterial transformation**

5 µL DNA was transformed in to 50 µL Top10 *E.coli* and incubated on ice for 30 minutes followed by heat shock at 42 °C for 35 seconds. 250 µL SOC media was added and incubated for 30 minutes at 37 °C, 250 rpm. Cultures were inoculated on agar plates containing 50 µg/mL ampicillin overnight at 37 °C. Colonies were extracted and added to 10 mL LB broth containing 50 µg/mL ampicillin overnight at 37 °C, 250 rpm. Plasmids were purified using QIAGEN QIAprep Spin Miniprep kit and Sanger sequenced to confirm the presence of the incorporated mutation.

2.3.7. **Plasmid DNA purification**

DNA was purified using the QIAprep Spin Miniprep kit following the manufacturers protocol. Briefly, the cell pellet was re-solubilised in 500 µL of buffer P1 followed by the addition of 500 µL buffer P2 and mixing by inversion then the addition of 700 µL N3. Samples were centrifuged at 15,000 *xg* for 10 minutes and supernatant remove and applied to a QIAprep Spin column and centrifuged at 150,000 *xg* for 1 minute to collect the DNA. The columns were washed sequentially with 500 µL buffer PB and buffer PE and centrifuged at 15,000 *xg* for 1 minute after each. DNA was eluted from the column by the addition of nuclease free H_2O and DNA concentration determined by nanodrop.

2.4. Cell-based sample preparation

2.4.1. **Cell culture**

U2OS T-Rex Flp-in cells were maintained in DMEM supplemented with 10% (v/v) fetal bovine serum, penicillin (100 U/mL), streptomycin (100 U/mL) and L-glutamine (2 mM), at 37 °C, 5% CO_2. At 80 % confluence, cells were washed with PBS, released with trypsin (0.05 % (v/v)) and centrifuged at 220 x *g.*

2.4.2. **Cell transfection**

To generate stable transfected cells, 1 μg plasmid (FLAG or MYC- WT & G95R PLK4) was incubated with 1 μL lipofectamine and made up to 500 μL with serum free media. Plasmids were co-transfected with a Flp recombinase, pOG44, of which 9 μg was incubated with 9 μL lipfoctamine 2000 (Invitrogen) in a 500 μL volume made up with serum free media. The 1:1 lipofectamine:plasmid mixture was added to the cells. Media was changed 24 hours later and after a further 24 hours, media was supplemented with hygromycin.

For transient transfection of MYC-WT PLK4, the above procedure was followed with the following exceptions: 6 μg of plasmid was transfected with 6 μL lipofectamine 2000 and pOG44 was omitted to prevent incorporation of the plasmid in to the genome.

2.4.3. **SILAC labelling**

U2OS T-Rex Flp-in cells stably transfected with FLAG-WT PLK4 or FLAG-G95R PLK4 were grown in DMEM supplemented with 10% (v/v) dialysed fetal bovine serum, penicillin (100 U/mL) and streptomycin (100 U/mL). Once 80 % confluency was reached, cells were split in to DMEM containing heavy labelled, ^{13}C ^{15}N labelled arginine and lysine for ~seven cell doublings to permit full incorporation of the label. At each passage, an aliquot of cells were removed and analysed by LC-MS/MS (see below) to assess label incorporation. At 80 % confluence, cells were washed with PBS, released with trypsin (0.05 % (v/v)) and centrifuged at 220 x *g.*

2.4.4. **Cell Lysis**

Cells were lysed with two different lysis buffers depending on whether western blot or MS analysis was required. For western blot analysis, cells were lysed with 1 % (v/v) NP-40, 1 % SDS in 50 mM Tris pH 8.0, 150 mM NaCL containing with 1x protease inhibitor cocktail tablet (Roche) (for western blot analysis). The lysate was sonicated briefly and centrifuged at 15,000 *x g* at 4 °C for 20 minutes. Protein concentration was determined using the Bradford assay.

To preserve protein-protein interactions, cells used for co-immunoprecipitation (IP) experiments were lysed with 50 mM Tris pH 8.0, 150 mM NaCl, 0.5 % NP-40, 1 mM DTT, 2

mM $MgCl_2$, benzonase supplemented with 1x protease inhibitor (Roche) cocktail tablet. Cells were incubated with lysis buffer for 30 minutes and centrifuged at 15,000 *x g* at 4 °C for 20 minutes. Protein concentration was determined using the Bradford assay.

For MS analysis, cells were lysed with an MS compatible buffer (0.25% (v/v) Rapigest SF (Waters) in 50 mM ammonium bicarbonate, supplemented with 1x PhosStop inhibitor (Roche) and sonicated (3x 10 seconds on, 1 minute off) to shear DNA. Samples were centrifuged at 15,000 *x g* at 4 °C for 20 minutes. Protein concentration was determined using the Bradford assay. SILAC labelled protein lysates (1 µg) were analysed by LC-MS/MS (see below) to assess label incorporation and subsequently mixed 1:1.

2.4.5. **Immunofluorescence Microscopy**

U2OS MYC-WT PLK4 or MYC-G95R PLK4 stably transfected cells were split in to a 6 well plate containing cover slips. At 80 % confluence, 1 µg/mL tetracycline was added to induce protein expression for 18 hours. Cells were washed with 2.5 mL PBS and fixed for 20 minutes with 2.5 mL of 3.7% paraformaldehyde. Cells were washed 3x with 2.5 mL PBS and permeabilised with 0.1 % (v/v) triton x-100 for 10 minutes. Cells were washed a further 3x with PBS and blocked with 1% (w/v) BSA, 0.3 M glycine for 60 minutes and washed 2x with PBS. Primary antibodies, anti-9E10 anti-myc antibody (1/100), anti-γ- tubulin (1/5000) or anti-pericentrin in 500 µL of 1% BSA were added to each coverslip and incubated overnight at 4 °C in a humidified chamber. Cells were washed 5x 10 minutes with PBS and secondary antibodies, rabbit anti mouse-Cy3 and goat anti rabbit-FITC (1/200) were added to each coverslip and incubated for 1 hour, room temperature in the dark. Cells were washed 5x 5 minutes with PBS and incubated with 1 µg/mL DAPI for 3 minutes. Coverslips were mounted on to microscope slides with immune-mount and imaged at the Centre for Cell Imaging (Institute for Integrative Biology) on a Zeiss LSM 880 confocal microscope using an alpha Plan-Apochromat 100x/1.46 DIC grade oil immersion lens. A 405 nm diode, 488 nm Argon and 561 nm diode laser were used to excite DAPI, Alexa 488 and Cy3 fluors respectively with a transmitted light image also captured. A pinhole of ~1 Airy unit was maintained throughout imaging. Spatial sampling was conducted at 2x Nyquist sampling for the chosen magnification with four times line averaging used to reduce noise. For each condition at least one z-stack through the sample was acquired with slices taken at ~1 µm throughout the cell of interest.

2.4.6. **Flow Cytometry for DNA analysis (FACS)**

U2OS FLAG-WT & G95R PLK4 cells grown in DMEM media were split in to 10cm^2 dishes. At 60 % confluence, cells were incubated with 1 µg/mL tetracycline for 18 hours to induce protein expression before incubating with 300 nM centrinone for 4 hours. Cells were washed with PBS and released with trypsin (0.05 % (v/v)). Following centrifugation at 220 xg, PBS was removed and cells were fixed with 70 % (v/v) ethanol, added slowly whilst vortexing to avoid aggregation and stored at -20 °C overnight. Fixed cells were washed with PBS and incubated with 200 µL Guava cell cycle reagent for 30 minutes in the dark at room temperature. Samples were then transferred to a 96 well plate and analysed using a Guava easycyte HT cytometer. ANOVA statistical analysis and Tukey post-hoc testing was performed in R.

2.5. Protein purification

2.5.1. **Expression of PLK4 & Aurora A plasmids**

pET30 N-terminal 6His- tagged WT or G95R PLK4 (1-269) plasmids were transformed in to Rosetta *E.coli* as described above (2.3.6). Cultures were plated on agar plates containing 50 µg/mL kanamycin and 37 µg/mL chloramphenicol overnight at 37 °C. Colonies were extracted and added to 100 mL LB broth containing 50 µg/mL kanamycin and 37 µg/mL chloramphenicol overnight at 37 °C, 250 rpm. Cultures were transferred to two 2L flasks and incubated at 37 °C, 250 rpm until OD_{600}= 0.6 – 0.8 was reached. Protein expression was induced upon the addition of 0.4 mM IPTG and incubated overnight at 18 °C, 250 rpm. Bacterial cells were harvested by centrifugation at 5000 *x g* for 10 minutes.

To generate dephosphorylated WT Aurora A, N-terminal pET30 6His-tagged Aurora A plasmids was co-transformed with pET30 6His-GST-tagged λ-phosphatase, following the procedure described above.

2.5.2. **Bacterial Cell Lysis**

Bacterial cells were lysed with 20 mL lysis buffer (50 mM Tris pH 7.4, 300 mM NaCl, 1 mM DTT, 10 % glycerol, 1 % Triton x-100, 10 mM imidazole, 0.1 mM EDTA, 0.1 mM EGTA,

supplemented with 1x protease inhibitor (Roche)) and sonicated to shear DNA. Samples were centrifuged at 14,000 *xg* for 1 hour and 8 °C to pellet insoluble material. Supernatant was filtered with a 0.22 µm filter prior to protein purification.

2.5.3. **Ni^{2+} - affinity purification**

Ni^{2+} - affinity purification was performed for all expressed proteins used in this thesis. Ni-NTA affinity resin (Thermo Fisher Scientific) (1 mL) was applied to a gravity flow column, washed with milliQ H_2O and equilibrated with 25 mL wash buffer (50 mM Tris pH 7.4, 500 mM NaCl, 10 % glycerol, 20 mM imidazole, 1 mM DTT). The filtered supernatant was then applied to the column; the flow through collected and the column was washed with a further 25 mL wash buffer. Bound protein was then eluted with increased imidazole concentration (500 mM). Elution buffer was added in 2 mL volumes and 500 µL fractions were manually collected.

2.5.4. **Glutathione-S-Transferase (GST) affinity purification**

To separate 6His-tagged Aurora A plasmids from 6His- GST-tagged λ-phosphatase, samples were incubated with 1 mL GST beads (equilibrated with wash buffer (50 mM Tris pH 7.4, 500 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol, 1 mM DTT) overnight at 8 °C. Proteins were eluted with 50 mM Tris-HCl pH 7.4, 10 mM glutathione and 5 mM DTT and 500 µL fractions were manually collected.

2.5.5. **Gel filtration chromatography**

6His-tagged WT & G95R PLK4 and 6His-tagged WT Aurora were purified by gel filtration chromatography, following Ni^{2+} - affinity purification. Samples were loaded on to an equilibrated (50 mM Tris pH 7.4, 100 mM NaCl, 10% (v/v) glycerol, 1 mM DTT) Superdex 16 600 column attached to an AKTA FPLC system. Flow rate was set at 1.5 mL/min and fractions (500 µL) were collected across the run (~2 hours).

2.5.6. **Strong Cation Exchange Chromatography**

Strong cation exchange chromatography was employed to 6His-tagged Aurora A plasmids from 6His- GST-tagged λ-phosphatase based on differences in PI between the two proteins (pI of Aurora A = 9.5; pI of λ-phosphatase= 4.5). Samples were diluted in loading buffer (50 mM Tris pH 7.4, 100 mM NaCl, 10% (v/v) glycerol, 1 mM DTT) and loaded on to a HiTrap SP HP cation exchange column at 1 mL/min flow rate, with chromatography controlled by an AKTA FPLC system. Bound protein was eluted using an increasing salt gradient from 0 - 100 % elution buffer (50 mM Tris HCl, pH 7.4, 1 M NaCl, 10% (v/v) glycerol, 1 mM DTT) and collected in 500 µL fractions.

2.6. Analytical methods

2.6.1. **Bradford assay**

Protein concentrations were determined using the Bradford assay. A standard curve was generated using BSA standards of known concentration (0-1 mg/mL). Protein samples were typically diluted with the appropriate buffer and 5 µL sample (or 5 µL buffer) added to 200 µL Bradford reagent and incubated for 5 minutes. The absorbance was read using a spectrophotometer at 595 nm and protein concentration determined using the calibration curve.

2.6.2. **Immunoprecipitation**

Anti-FLAG M2 agarose resin (30 µL) was washed 3x with 50 mM Tris pH 8.0, 150 mM NaCl wash buffer and incubated with cell lysate overnight at 4 °C. Samples were centrifuged at 1000 rpm for 1 minute and flow through removed. Beads were washed 5 x with 50 mM Tris pH 8.0, 150 mM NaCl wash buffer and proteins were eluted from the antibody by boiling with 2x SDS sample buffer.

2.6.3. **SDS-PAGE**

Proteins were analysed by SDS-PAGE under reducing conditions using 8 % polyacrylamide gels for analysis of full length PLK4 (MW- 110 kDa) or 10 % polyacrylamide gels for analysis of recombinant proteins (PLK4 1-269, MW- 32 kDa). Samples were centrifuged at 15,000 *x g* for 15 minutes prior to the addition of 5x SDS sample buffer (0.25 M Tris-HCl pH 6.8, 10 % SDS, 50 % glycerol, 0.5 M DTT, 0.25 % bromophenol blue) and boiled at 95 °C for 5 minutes. Samples were loaded on to the gel and run at 200 V for approximately 45 minutes. Protein bands were visualized by Coomassie Brilliant Blue staining (0.1 % Coomassie G250, 45 % methanol, 10 % acetic acid).

2.6.4. **Western blot**

Proteins were electrotransferred from the gel on to a nitrocellulose membrane at 100 V for 60 minutes. The membranes were incubated in blocking buffer (5 % (w/v) milk powder in TBST) for 1 hour and washed 3x 5 minutes with TBST. Primary antibodies were incubated with the membrane overnight at 4 °C. Membranes were washed 3 x 10 minutes with TBST and incubated with secondary antibody for 1 hour at room temperature. The antibody-antigen complex was visualised upon addition of ECL substrate and exposing film to the membrane.

2.7. *In vitro* kinase assays

2.7.1. **Differential Scanning Fluorimetry (DSF)**

Proteins were diluted with 50 mM Tris pH 8.0, 150 mM NaCl to a final concentration of 5 µM and incubated with 1 mM ATP +/- 10 mM $MgCl_2$ or 10 mM centrinone/VX-680. Samples were incubated at room temperature for 15 minutes before the addition of SYPRO Orange fluorescent dye (Invitrogen). The samples were analysed using an Applied Biosystems StepOnePlus Real-Time PCR instrument using a DSF thermal-shift protocol developed for the analysis of protein kinases (Rudolf *et al.*, 2014). The temperature was increased in 0.3 °C intervals from 25 °C to 95 °C. DSF assays were performed with three technical replicates.

2.8. Enzymatic assays

2.8.1. **PLK4 activity**

The Caliper LabChip EZ Reader platform measures enzyme activity by assessing the mobility shift of a fluorescently labeled peptide substrate, which changes upon phosphorylation and can be quantified by comparative integration of phosphorylated and dephosphorylated peptide peaks. To assess WT and G95R PLK4 activity, increasing amounts of PLK4 (0.5 µg to 10 µg) were incubated with 1 mM ATP (to mimic cellular concentration) and 5 µM of a fluorescent kemptide-derived peptide substrate (5-FAM-FLAKSFG**S**PNRAYKK) in 25 mM Hepes, 10 mM $MgCl_2$, and 0.001% (vol/vol) Brij 35 buffer. Reactions were pre-incubated for 30 minutes at 37 °C and measurements were taken over 4 EZ Reader cycles at room temperature.

2.8.2. **PLK4 inhibition**

To assess WT and G95R PLK4 inhibition with centrinone, 5 µg of protein was incubated with increasing half-log concentrations of centrinone (10 nM to 100 µM), 1 mM ATP (to mimic cellular concentration) and 5 µM of the fluorescent kemptide-derived peptide substrate (5-FAM-FLAKSFG**S**PNRAYKK) in 25 mM Hepes, 10 mM $MgCl_2$, and 0.001% (vol/vol) Brij 35 buffer. Reactions were pre-incubated for 30 minutes at 37 °C and measurements were taken over 4 EZ Reader cycles at room temperature.

2.8.3. **Aurora A inhibition**

To assess WT and G216L Aurora A inhibition with centrinone, 10 ng of protein was incubated with increasing concentrations of centrinone (10 nM, 100 nM, 1000 nM) and 2 µM of the fluorescent kemptide-derived peptide substrate (5FAM-LRRSLG) in 25 mM Hepes, 10 mM $MgCl_2$, and 0.001% (vol/vol) Brij 35 buffer. Reactions were pre-incubated for 15 minutes at 37 °C and measurements were taken over 4 EZ Reader cycles at room temperature.

2.8.4. **PLK4 : Aurora A activation**

To assess the effects of GST, GST-TPX2, WT PLK4, or kinase-dead (D154A) PLK4 on inactive Aurora A, a 10-fold molar excess of each protein was incubated with 100 ng λ-phosphatase-treated Aurora A with 1 mM ATP (to mimic the cellular concentration) and 10 mM $MgCl_2$ for 30 min at 30°C. Aurora A activity was then assessed in kinetic mode by the addition of a fluorescent kemptide-derived peptide substrate (5FAM-LRRSLG), with calculation of peptide phosphorylation every minute for 30 cycles at 20°C (1 min, approximately one cycle). To measure enhanced effects of WT PLK4 (1–285) on the activity of catalytically active (non–phosphatase treated) Aurora A, 2 µM of the Aurora A fluorescent kemptide-derived peptide substrate was incubated with 30 ng PLK4 and 10 ng Aurora A in 25 mM Hepes, 10 mM MgCl2, and 0.001% (vol/vol) Brij 35 buffer, with 1 mM ATP. Reactions were preincubated for 15 min at 37 °C to permit activation of Aurora A, and measurements were taken over 60 EZ Reader cycles at room temperature in kinetic mode.

2.8.5. ***In vitro*** **kinase activation assay**

Purified "kinase-dead" N-terminally 6His-tagged D274N Aurora A (available in the laboratory) (1.5 µg) was incubated with 3 µg of 6His-tagged WT Plk4 (1–285) lacking the polo box domains or a catalytically inactive D154A Plk4 (1–264) mutant (±10 µM centrinone) at 37°C in 50 mM Tris, pH 7.4, 10 mM $MgCl_2$, 1 mM DTT, and 1 mM ATP. Aliquots were removed at 0, 10, 30, 60, 120, and 240 min after reaction initiation and stopped by boiling in SDS sample buffer. To evaluate site-specific Aurora A phosphorylation, trypsin proteolysis and liquid chromatography–tandem mass spectrometry (LC-MS/MS) analysis was performed at the 120-min time point.

2.9. Sample preparation for LC-MS

2.9.1. **Trypsin digestion**

Proteins were digested in to tryptic peptides as follows: disulfide bonds were reduced by addition of 3 mM DTT in 50 mM ammonium bicarbonate and heated at 60°C for 15 minutes. The resulting free cysteine residues were alkylated with 14 mM iodoacetamide (dark, room temperature, 45 minutes) and excess iodoacetamide quenched by addition of DTT to a final concentration of 7 mM. Proteins were digested overnight with trypsin (2% (w/w); Promega) at 37 °C. RapiGest SF hydrolysis was induced by addition of trifluoroacetic acid (TFA) to 1 % (v/v) and incubated at 37 °C for up to 2 h, 400 rpm. Insoluble hydrolysis product was removed by centrifugation (13,000 x *g*, 15 min, 4°C).

2.9.2. **Desalting**

Peptides generated from U2OS cell lysates were desalted using C18 macro columns (Harvard Apparatus, Cambridge, UK). Briefly, columns were conditioned with 100 % methanol and washed with H_2O and 1% (v/v) TFA. Peptides were loaded on to the column and centrifuged for 1 minute at 110 x *g*. The flow-through was re-applied a total of 5 times and peptides were eluted with 80% (v/v) MeCN and 1% (v/v) TFA and dried to completion by vacuum centrifugation.

2.9.3. **High pH reversed phase fractionation**

Tryptic peptides generated from SILAC labelled protein lysates were dissolved in 94.5 % buffer A (20 mM ammonium hydroxide pH 10) and 5.5 % buffer B (20 mM ammonium hydroxide in 90% acetonitrile, 10% water) and loaded on to a Extend C18 (3.5 µm, 3mm x 150 mm) Agilent column. Peptides were eluted following an increase in buffer B up to 30% over 25 minutes then 75% B over 12 minutes at a flow rate of 0.5 mL/minute. Sixty 500µL fractions were collected, partially dried by vacuum centrifugation and concatenated in to twelve pools. Aliquots (5 µg) were removed from each pool for total proteomics analysis. Samples were dried to completion by vacuum centrifugation.

2.9.4. **Titanium dioxide enrichment**

Dried peptides were dissolved in loading buffer (80 % (v/v) MeCN, 5 % (v/v) TFA, 1 M glycolic acid), sonicated and incubated with 5 mg titanium dioxide resin (5:1 (w/w) beads:protein; GL Sciences) at 1400 rpm for 10 minutes on a thermomixer. Wash steps were performed sequentially with 150 µL loading buffer; 150 µL wash buffer 1 (80% (v/v) MeCN, 1% (v/v) TFA) and 150 µL wash buffer 2 (10 % (v/v) MeCN, 0.2% (v/v) TFA). Phosphopeptides were eluted with increasing pH (1 % (v/v) ammonium hydroxide and 5 % (v/v) ammonium hydroxide) and dried to completion by vacuum centrifugation. Peptides were re-solubilized in 240 µL of 96 % (v/v) H_2O, 3 % (v/v) MeCN, 1 % (v/v) TFA.

2.9.5. **Preparation of phosphopeptide library**

The phosphopeptide library (Intavis) containing 180 chemically synthesized phosphorylated peptides (~10 pmol) were reconstituted in 10 µL H_2O. The phosphopeptides were split in to five pools, designed to separate phosphoisomers and thus ensure confidence in phosphosite localization.

2.10. Mass Spectrometry

2.10.1. **Thermo Orbitrap Fusion**

Reversed-phase capillary HPLC separations were performed using an UltiMate 3000 nano system (Dionex) coupled in-line with a Thermo Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific, Bremen, Germany). Samples were loaded onto the trapping column (PepMap100, C18, 300 µm x 5 mm), using partial loop injection, for seven minutes at a flow rate of 9 µL/min with 2% (v/v) MeCN, 0.1% (v/v) TFA and then resolved on an analytical column (Easy-Spray C18 75 µm x 500 mm, 2 µm bead diameter column) using a gradient of 96.2% A (0.1% (v/v) formic acid (FA)): 3.8% B (80% (v/v) MeCN, 0.1% (v/v) FA) to 50% B over 97 minutes at a flow rate of 300 nL/min.

MS(/MS) data were acquired on an Orbitrap Fusion as follows: all MS1 spectra were acquired over *m/z* 350-2000 in the orbitrap (120K resolution at 200 m/z for high-low strategies and 60K resolution at 200 m/z for high-high strategies); automatic gain control

(AGC) was set to accumulate $2x10^5$ ions, with a maximum injection time of 50 ms. Data-dependent tandem MS analysis was performed using a top speed approach (cycle time of 3 s) with multiple fragmentation methods tested (see Appendix, Supplementary Table S1). The normalized collision energy was optimized at 32% for HCD. MS2 spectra were acquired with a fixed first *m/z* of 100. The intensity threshold for fragmentation was set to 50,000 for orbitrap methods and 5000 for ion trap methods and included charge states 2+ to 5+. A dynamic exclusion of 60 seconds was applied with a mass tolerance of 10 ppm. For neutral loss triggered ETcaD/EThcD methods, fragmentation was enabled for all precursor ions exhibiting neutral loss of mass 97.9763 Da or 80 Da with a mass tolerance of 20 ppm for orbitrap data and 0.5 *m/z* for ion trap data, where the neutral loss ion was one of the top 10 most intense MS2 ions. ETD calibrated parameters were applied. AGC was set to 10,000 with a maximum injection time set at 50 ms for IT and 70 ms for OT; ETD reaction time was charge-dependent.

2.11. Data Analysis

2.11.1. **Proteome Discoverer 1.4**

Data were processed using Thermo Proteome Discoverer (v. 1.4) in conjunction with MASCOT (v 2.6). To address the requirement of MASCOT for centroided data, raw data files were converted to mzML format in order to perform MS2 de-isotoping prior to processing with MASCOT through the Proteome Discoverer (PD) pipeline. Peak lists were searched against a database containing either the synthetic phosphopeptide sequences or the human UniProt database (downloaded December 2015; 20187 sequences). Parameters were set as follows: MS1 tolerance of 10 ppm; MS2 mass tolerance of 0.01 Da for orbitrap detection, 0.6 Da for ion trap detection; enzyme specificity was set as trypsin with 2 missed cleavages allowed; no enzyme was defined for the phosphopeptide library processing; carbamidomethylation of cysteine was set as a fixed modification; phosphorylation of serine, threonine and tyrosine, and oxidation of methionine were set as variable modifications. Non-fragment filtering was applied to ETD scans to remove the precursor peak within a 4 Da window and remove charged reduced precursor and neutral loss ions from charged reduced precursor ions within a 2 Da window. *ptm*RS was run in PhosphoRS mode using diagnostic fragment ions and analyzer specific fragment ion tolerances as

previously defined in the search. For EThcD data, 'Treat all spectra as EThcD' option was set to 'True'. Data was filtered to a 1% false discovery rate (FDR) on PSMs using automatic decoy searching with Mascot and a target-decoy search with Andromeda.

For processing of SILAC data, ^{15}N ^{13}C Arginine and ^{15}N ^{13}C Lysine were set as a variable modifications.

2.11.2. **MaxQuant**

Phosphopeptide data from the MS acquisition method optimization study were processed using Andromeda with PTM-score implemented within MaxQuant (version 1.6.0.16) (Cox & Mann, 2008). MS/MS spectra were searched against a database containing the human UniProt database (downloaded December 2015; 20187 sequences). Trypsin was set as the digestion enzyme and two missed cleavages were permitted. Cysteine carbamidomethylation was set as a fixed modification. Variable modifications were set as oxidation (M), phospho (STY). Default instrument parameters and score thresholds were used: MS1 first search peptide tolerance of 20 ppm and main search tolerance of 4.5 ppm; FTMS MS2 tolerance of 20 ppm; ITMS MS2 tolerance of 0.5 Da. A false discovery rate of 1 % for peptide spectrum matches (PSM) and proteins was applied.

For processing of SILAC data, Thermo raw. files were loaded in to MaxQuant (version 1.6.0.16). Total proteomics (protein expression data) and phosphoproteomics experiments were processed separately. The experimental template design separated individual bioreplicates in to separate experiments, linking the experiment to the relevant fractions. MaxQuant parameters were as described above, with the following additions: ‘multiplicity’ was set to 2 and Arg10, Lys8 were selected as labels. For proteomics datasets,the variable modifications oxidation (M), acetyl (protein N-term) were included. Both ‘requantify’ and ‘match between runs’ were enabled. Two peptides were required for protein quantification.

2.11.3. **Perseus**

Post processing was performed using Perseus (version 1.6.0.7). For proteingroups.txt output files, Perseus was used to filter out contaminants, reverse decoy hits and those ‘matched only by site’. Additionally, data were filtered to include proteins identified in ≥3

(out of 4) bioreplicates and Log2 transformed. For phosphosites(STY).txt output files, data were filtered as above. In addition, 'expand site table' feature was used to separate individual phosphosites, and a phosphosites localisation cut-off of ≥0.75 was applied. Ratios were Log2 transformed. Statistical analysis was then performed using the LIMMA package in R with Benjamini-Hochberg multiple corrections to generate adj. *p* values (performed by Dr. Simon Perkins).

2.11.4. **Arginine to proline conversion**

Non-normalised SILAC ratios were used to assess arginine to proline conversion. The peptides.txt MaxQuant output files was loaded in to Perseus (version 1.6.0.7) and reverse decoy hits and contaminants were removed. Non-normalised SILAC ratios were Log2 transformed and the density of the SILAC ratios against the proline count observed in the identified peptides was calculated. Log2 SILAC ratios were plotted against proline count.

2.11.5. **Manual spectra annotation**

MS2 spectra were extracted from the Thermo raw. files and manually annotated using the MASCOT output and Protein Prospector (version 5.22.1). Spectra were redrawn in R to improve figure quality.

2.11.6. **DAVID**

Functional annotation and enrichment analysis of regulated proteins and phosphosites was performed using DAVID (version 6.8) (Huang da *et al.*, 2009). Information pertaining to enriched cell compartments, molecular functions, biological processes and KEGG pathways were exported. Results were filtered to exclude a protein count <10 and redundant GO terms. Benjamini-Hochberg adjusted *p* values were used to determine significance and terms with *p* value <0.05 were highlighted to indicate significance. Unfiltered tables exported from DAVID can be found in Appendix 3, Table 9.6 to 9.9.

2.11.7. **STRING**

Significantly downregulated proteins from the G95R PLK4 SILAC dataset (adj. *p* value <0.05) were submitted to the STRING database (version 10.5) to assess protein-protein interactions (Szklarczyk *et al.*, 2015). A high confidence (score 0.7) filter was applied, and only 'experiments' and 'databases' as active interaction sources were included.

2.11.8. **IceLogo**

Sequences surrounding Ser phosphorylated peptides identified as downregulated were extracted from the 'sequence window' column in the MaxQuant phosphoSTY.txt output and input into IceLogo (Colaert *et al.*, 2009). *Homo sapien* Swiss-Prot protein sequence database was used as the reference dataset. A P value threshold (*p*= 0.01) was applied to determine motif confidence.

3. Development of biochemical tools for the functional analysis of PLK4

3.1. Introduction

The aim of this work is to understand PLK4-dependent signalling pathways and to identify potential PLK4 substrates by phosphoproteomics. This chapter describes the development of biochemical tools for the functional analysis of PLK4 in human cells.

Despite the importance of PLK4 in regulating centrosome duplication, very few substrates have been identified, the majority of research having focused on the interactions between PLK4 and the centriolar proteins STIL and SAS6, which are involved in centriole duplication (Dzhindzhev *et al.*, 2014; Ohta *et al.*, 2014). Despite advances in the understanding of how these proteins co-ordinate their activities to regulate centriole duplication, gaps remain in the understanding of centriole formation and additional PLK4 substrates have yet to be identified. In addition, a wider role for PLK4 and its implications in cell signalling pathways has not been explored thus far.

Table 3.1 Known PLK4 substrates. A list of direct PLK4 substrates reported in the literature to date and whether the interaction was determined by *in vitro* and/or *in vivo* based methods.

PLK4 substrate	Evidence	Reference
STIL Ser1108, Ser1116	*In vitro/in vivo*	(Arquint *et al.*, 2015); (Dzhindzhev *et al.*, 2014); (Ohta *et al.*, 2014); (Kratz *et al.*, 2015); (Moyer *et al.*, 2015).
CP110 Ser98	*In vitro/in vivo*	*(Lee et al., 2017)*
PCM1 Ser372	*In vitro/in vivo*	*(Hori et al., 2016)*
GCP6 Ser7, Ser15, Ser392, Ser1437, Ser1465	*In vitro*	(Bahtz *et al.*, 2012)
Cep152/Cep192 Phosphosites unknown	*In vitro*	(Hatch *et al.*, 2010); (Sonnen *et al.*, 2013)
FBXW5 Ser151	*In vitro*	(Puklowski *et al.*, 2011)

The levels of PLK4 in the cell are kept very low through tight regulation of protein abundance via autophosphorylation and proteasomal degradation. Utilising a selected reaction monitoring (SRM) mass spectrometry approach, PLK4 was determined to be expressed at less than ~4000 copies per cell. This is in comparison to other centrosomal proteins such as SAS6 which is expressed at ~20,000 copies per cell (Bauer *et al.*, 2016). The low expression of PLK4 in human cells therefore makes it difficult to study endogenous protein activity, without chemical arrest of the cycle cycle (Leung *et al.*, 2002). There is, therefore, a need to establish a suitable over-expression system in which protein expression can be carefully controlled to ensure cellular conditions mimic those of the rapidly turned over endogenous protein as closely as possible (Habedanck *et al.*, 2005). The first part of this chapter will discuss the cloning and generation of PLK4 plasmids, their transfection into a Tet-inducible U2OS cell line and the analysis of PLK4 protein expression and subcellular localisation.

The human bone osteosarcoma U2OS T-Rex cell line was chosen for stable transfection of the PLK4 plasmids due to its unique phenotype. U2OS cells contain highly altered chromosome numbers, often in the hypertriploid range. However, unlike many other cancerous cell lines, the cells do not undergo multipolar mitosis due to functional activities of the tumour suppressor proteins p53 and pRB. Consequently, the control of centrosome number is maintained and the normal duplication process can be studied (Niforou *et al.*, 2008). A further advantage of this cell system is conferred by the use of a stable, rather than transient, transfection procedure that allows for controlled protein expression at levels much lower than that obtained by transient transfection. Transient over-expression of PLK4 leads to aberrant activity resulting in over-duplication of the centrioles and formation of extra centrosomes (Bettencourt-Dias *et al.*, 2005; Habedanck *et al.*, 2005). Expression of the transfected gene is controlled by a $TetO_2$ sequence adjacent to the promoter. In the absence of tetracycline, the $TetO_2$ sequence binds molecules of a Tet repressor that prevent transcription of the gene. Upon addition of tetracycline, the repression at the $TetO_2$ sites is relieved through the binding of tetracycline to the Tet repressors. Controlling protein expression through the use of a stable, Tetracycline-inducible cell line allows the conditions to closely resemble that of the endogenous protein with respect to localisation and activity, with a high enough level of expression to study by phosphoproteomics.

Phosphoproteomics experiments to study protein kinase-based signalling typically utilise small molecule inhibitors that can inhibit the activity of the protein of interest, so that changes in protein expression or phosphorylation stoichiometry can be assessed. However, a key issue with this approach is the known promiscuity of ATP-competitive small molecule inhibitors, due to the high degree of conservation of the ATP binding site in protein kinases (Hanks & Hunter, 1995; Bain *et al.*, 2007). This promiscuity can lead to 'off-target' effects and perturbations of signalling pathways that are not biologically relevant to the protein under study (Fabian *et al.*, 2005; Anastassiadis *et al.*, 2011; Klaeger *et al.*, 2017) Inhibitors with high levels of selectivity are continuously being synthesised and released to the research community and the most recently developed, highly specific PLK4 inhibitor, centrinone, was selected for the inhibition experiments described in this chapter. However, even a kinase inhibitor with high specificity will likely inhibit additional off-target protein kinases. Off-target inhibition greatly complicates downstream analysis and the ability to correctly attribute protein and phosphosite level changes to the activity of PLK4 directly and, therefore, additional tools to validate phosphoproteomics data are required.

A drug resistant mutant (G95R) that retains WT activity and localises correctly at the centrioles represents an ideal approach by which to separate on- and off-target effects of centrinone inhibition of PLK4. Gly95 is located at the '+6 position' adjacent to the hinge loop of PLK4 and is an important residue for sensitivity to kinase inhibitors such as VX-680 and MLN8054 (Bailey *et al.*, 2014). PLK1-3 contain an arginine residue at the +6 tetrad position and are not sensitive to VX-680-like classes of small molecule inhibitors (Johnson *et al.*, 2007). Mutation of Gly95 to Arg results in a catalytically-active protein resistant to inhibition by small molecule inhibitors within the ATP-binding site (Sloane *et al.*, 2010).

An additional challenge of a phosphoproteomics approach to studying PLK4 signalling is that in addition to PLK4, many other cell cycle proteins are also expressed at very low levels in the cell and are likely beyond the limit of detection of current mass spectrometers when used for high-throughput proteomics experiments (Choudhary & Mann, 2010; Wilhelm *et al.*, 2014). This is exacerbated for phosphopeptides by their often low stoichiometry and low ionization efficiency, which interferes with the identification of phosphopeptides from proteins of low abundance (Olsen & Mann, 2013; Solari *et al.*, 2015). Pre-fractionation of digested cell lysates is therefore crucial for reducing mass spectral sample complexity and increasing the analytical dynamic range allowing for greater coverage of the proteome.

High pH reversed phase fractionation has emerged as the leading method for maximising coverage of the phosphoproteome, with more than double the number of phosphopeptide identifications compared with the other common chromatography methods (Batth *et al.*, 2014). The development of a high pH reversed phase fractionation approach for maximising phosphopeptide identifications will also be discussed in this chapter.

As outlined in Chapter 1, stable isotope labelling of amino acids in culture (SILAC) is a commonly used labelling technique for quantitative phosphoproteomics studies. SILAC based quantification requires cells to be fully labelled prior to sample mixing, as partially labelled cells will contribute to the 'light' peptide signal upon analysis of the resulting peptides by MS, resulting in inaccurate quantification. It is; therefore, of critical importance to assess the labelling efficiency over a number of cell divisions to ensure that the heavy labels were fully incorporated into the U2OS cell line, which is described in this chapter.

3.1.1. **Aims**

The aim of the work described in Chapter 3 was to develop new tools for the proteomic analysis of PLK4 activity in human cells. These tools include FLAG-WT and G95R PLK4-expressing DNA plasmids for stable transfection in to U2OS cells, and the characterisation of protein expression and subcellular localisation to ensure the full-length proteins are being expressed and are correctly localised. In addition, this work aimed to establish the degree of similarity between the activities of WT & G95R PLK4, to confirm that G95R PLK4 expressing cells can be used as a tool for validating on- and off-target effects of centrinone in PLK4 phosphoproteomics studies. Finally, this work aimed to develop and implement stages of an experimental workflow to perform phosphoproteomics analysis, including offline fractionation by high pH reversed phase chromatography and assessment of the incorporation of heavy labelled arginine and lysine in SILAC experiments in U2OS cells.

3.2. Results & Discussion

3.2.1. **Generation of WT & G95R PLK4 plasmids**

To establish stably transfected T-Rex U2OS cell lines capable of controlled expression of PLK4, PCR was used to amplify full-length WT PLK4 cDNA (1-2913 bp), engineered to include a Kozak sequence (CCACCATG) required for high level initiation of translation of a mRNA in eukaryotic cells, followed immediately by an initiator 'ATG' codon. In addition, a FLAG peptide sequence (DYKDDDDK) was incorporated at the 5' end of the PLK4 cDNA to generate a small N-terminal tag that acts as an artificial antigen allowing for immunoprecipitation of the expressed protein (Figure 3.1A). The resulting DNA construct was then sub-cloned in to the pcDNA5 FRT/TO vector specific to the T-Rex cell system via the engineered *Not I* and *Hind III* restrictions sites (Figure 3.1B, C) and Sanger sequenced to confirm the sequence was correct. This pcDNA5 vector contains important features required for stable integration in to the host genome. These include a hybrid human cytomegalovirus/$TetO_2$ promoter for the controlled expression of the transfected protein in mammalian cells via tetracycline binding at this promoter; a FLP Recombination Target (FRT) site which binds FLP recombinase to initiate recombination of the vector in to the host genome and a hygromycin resistance gene permitting the selection of successfully transfected cells using this antibiotic (Figure 3.1B. C).

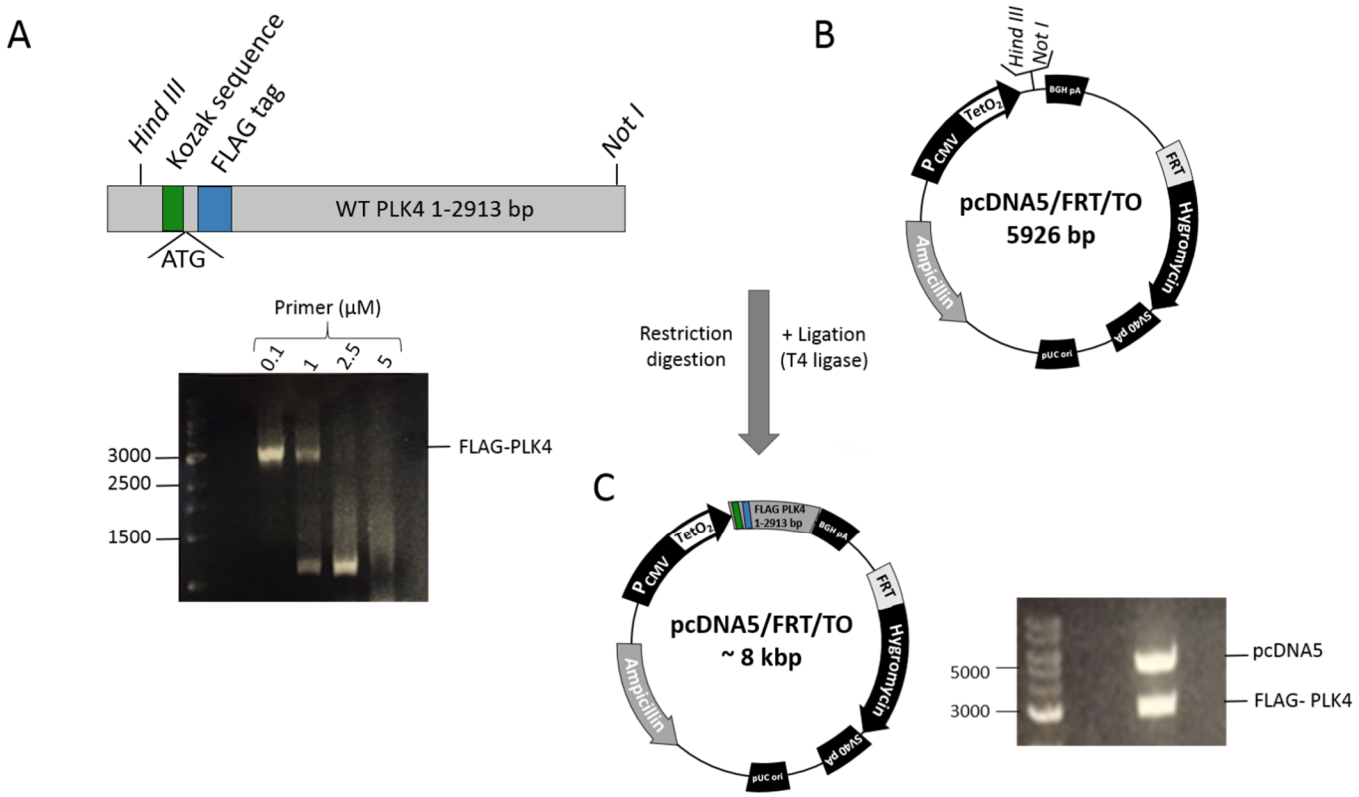


Figure 3.1. Cloning strategy to generate FLAG-PLK4 pcDNA5 constructs A) Full length wild type PLK4 cDNA was amplified by PCR to introduce unique restriction sites (Hind III & Not I), a Kozak sequence required for protein expression in eukaryotic cells and a FLAG tag for immunoprecipitation. A range of primer concentrations were tested and analysis of PCR product for the 0.1 µM and 1 µM primer conditions by agarose gel electrophoresis revealed a single band at ~3,000 bp B) PLK4 cDNA was cloned in to pcDNA5 vector using T4 ligase. C) To confirm successful ligation, plasmids were restriction digested and analysed by agarose gel electrophoresis, revealing a band at 3,000 bp and at 5,000 bp representing FLAG-PLK4 cDNA and pcDNA5 vector respectively**.**

Using the FLAG-PLK4 pcDNA5 FRT/TO construct, a drug-resistant PLK4 (G95R) mutant was generated by site-directed mutagenesis using overlapping primers incorporating the glycine to arginine mutation (Figure 3.2A). Sanger sequencing of the plasmid confirmed successful incorporation of the mutation (Figure 3.2B). Both the WT and G95R constructs could then be transfected and stably integrated in to the genome of T-Rex U2OS cells.

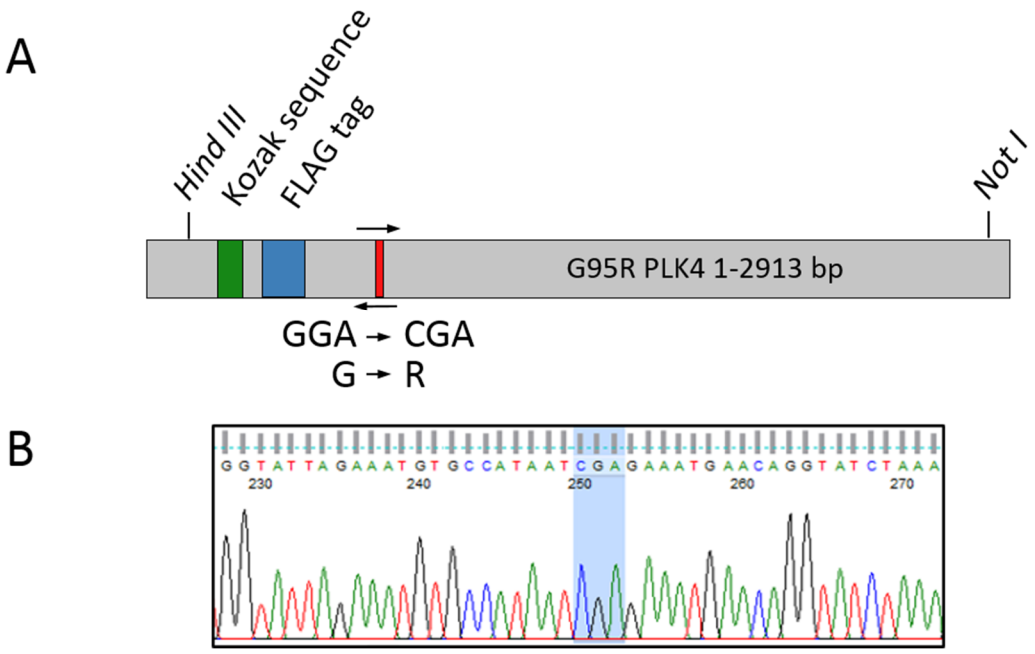


Figure 3.2. Generation of FLAG-G95R PLK4 by site-directed mutagenesis. A) Full length wild type FLAG PLK4 cDNA was mutated by site directed mutagenesis using forward and reverse primers containing a G to R mutation to produce drug resistant PLK4. B) Plasmids were sent for Sanger sequencing which confirmed successful incorporation of the mutation.

3.2.2. **Transection of plasmids into U2OS cells**

U2OS T-Rex Flp-In cells incorporate a pFRT/lacZeo plasmid consisting of a zeocin fusion gene under the control of an SV40 promoter and an FRT site downstream of an initiation codon. T-Rex cells can therefore be selected for (and the plasmid maintained) using DMEM supplemented with an appropriate concentration of zeocin (Figure 3.3A). FLAG-WT and G95R PLK4-pcDNA5 plasmids were transfected into T-Rex Flp-In U2OS cells (a kind gift from Dr Gopal Sapkota, MRC PPU, University of Dundee) using lipofectamine 2000 transfection reagent. The plasmids were co-transfected with a pOG44 plasmid encoding a FLP recombinase, which mediates homologous recombination (Figure 3.3B). This integrates FLAG-PLK4 between the FRT sites on the pFRT/lacZeo and pcDNA5 plasmids, which brings the SV40 promoter and initiation codon next to the hygromycin resistance gene present in pcDNA5 and inactivates the zeocin gene (Figure 3.3C). Successfully transfected cells were then selected for by supplementing media with hygromycin as they will have acquired hygromycin resistance, whereas unsuccessfully transfected cells will remain hygromycin sensitive and fail to establish. Successfully transfected cells were propagated to produce stably transfected, isogenic cell lines.

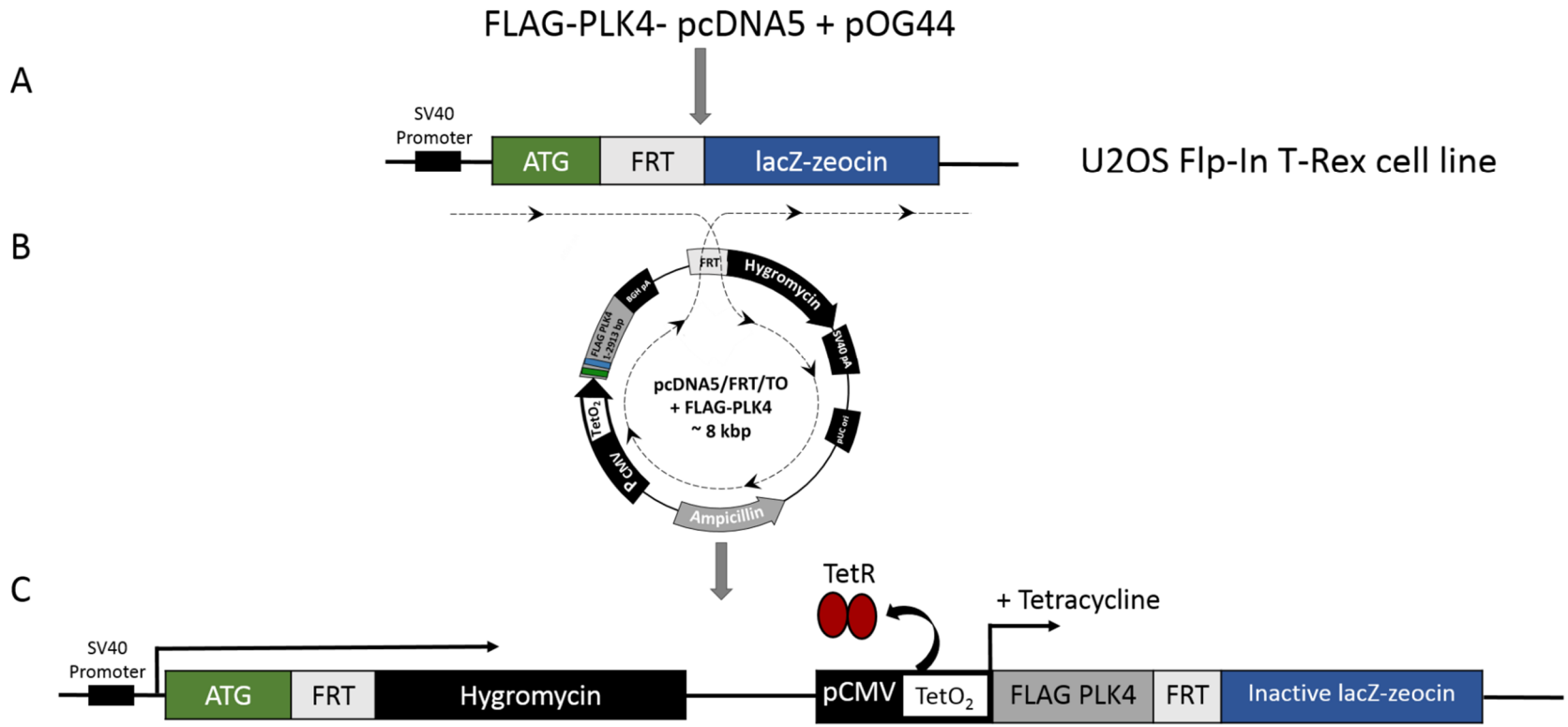


Figure 3.3. Summary of the U2OS T-Rex Flp-In system. a) FLAG-PLK4 pcDNA5 plasmids were co-transfected with pOG44 in to U2OS T-Rex Flp-In cells. b) Flp recombinase is expressed from the pOG44 plasmid and catalyzes homologous recombination between the FRT sites on pcDNA5 and the U2OS cell line, which integrates FLAG-PLK4 pcDNA5 in to the host genome. c) Integration of FLAG-PLK4 pcDNA5 confers cell resistance to hygromycin, providing a means to select successfully transfected cells. Addition of tetracycline relieves repression at $TetO_2$ sequence and results in expression of FLAG-PLK4 in a controlled manner.

FLAG-WT and G95R PLK4 protein expression levels were assessed upon induction of protein expression with tetracycline. Cells were also treated with the proteasome inhibitor MG-132, to prevent degradation of low-abundance PLK4 in proteasomes. As discussed in Chapter 1, PLK4 trans-autophosphorylates at residues Ser293 and Ser297 within its phosphodegron and subsequently binds the SCF-Slimb/βTrCP-E3 ubiquitin ligase to signal the protein for degradation by the proteasome (Holland *et al.*, 2012; Cunha-Ferreira *et al.*, 2013; Klebba *et al.*, 2013). Therefore, blocking PLK4 degradation results in an accumulation of PLK4 in the cell, which can be observed by western blot. Figure 3.4 confirms that both full-length WT and G95R FLAG-PLK4 are expressed in U2OS cells; the presence of an immunoreactive band at ~110 kDa increases in abundance in the presence of MG-132. The presence of faint bands at ~110 kDa in the negative controls is likely due to "leaky" expression of the transfected genes, where expression of the gene occurs in the absence of tetracycline. This is supported by the detection of a slightly more intense band in the 'negative control' MG-132-exposed cells, which reflects the accumulation of this protein due to stabilisation after inhibition of the proteasome.

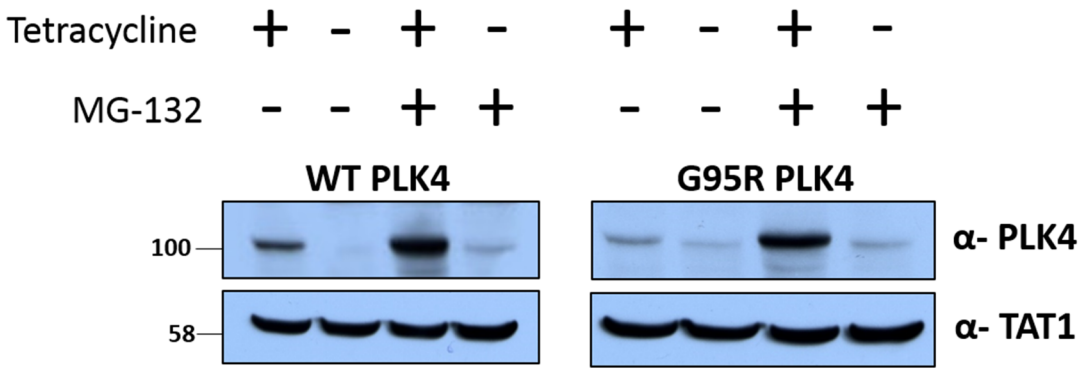


Figure 3.4. Overexpression of WT & G95R PLK4. U2OS cells stably transfected with WT or G95R PLK4 were incubated +/- 1 μg/mL tetracycline for 18 hours to induce protein expression, followed by +/- 10 μM MG-132 for 4 hours. Lysates were analysed by western blot and probed using the indicated antibodies. Full length PLK4 is revealed by the presence of a band at ~110 kDa, which increases in intensity following MG-132 treatment.

3.2.3. **Immunofluorescence**

To confirm that the over-expression of PLK4 was associated with targeting to the correct subcellular localisation, U2OS cells were imaged by immunofluorescence with a confocal microscope. With an established in-house protocol to study MYC-tagged centrosomal proteins by immunofluorescence, MYC- WT & G95R U2OS cells (data available in S.Ferries Masters report, 2015) were employed to study subcellular localisation. To confirm that the MYC- tagged proteins were expressed at a similar level to FLAG-PLK4, protein expression was induced in MYC-WT & MYC-G95R PLK4 or FLAG-WT & FLAG-G95R PLK4 by tetracycline, followed by treatment with MG-132. Figure 3.5 reveals the presence of a band at ~110 kDa in both the FLAG- and MYC- PLK4 cell lines, showing that both constructs produce full-length protein. In addition, the MYC-PLK4 proteins also accumulate within the cell when the proteasome is inhibited. Slightly increased levels of the MYC-PLK4 proteins can be observed compared with FLAG-PLK4 which further justified the use of the MYC cell lines for assessment of subcellular localisation.

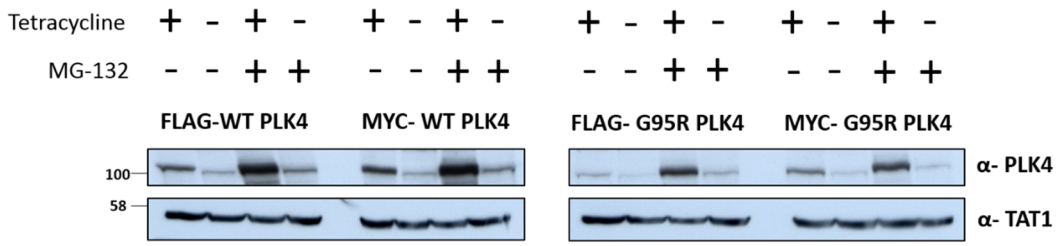


Figure 3.5. Overexpression of FLAG & MYC- WT & G95R PLK4. U2OS cells stably transfected with FLAG or MYC- WT or G95R PLK4 were incubated +/- 1 μg/mL tetracycline for 18 hours to induce protein expression, followed by +/- 10 μM MG-132 for 4 hours. Lysates were analysed by western blot and probed using the indicated antibodies. Full length PLK4 is revealed by the presence of a band at ~110 kDa, which increases in intensity following MG-132 treatment.

To establish unequivocal centrosomal localisation, pericentrin and γ-tubulin, known markers of the centrosome were imaged alongside WT & G95R PLK4. Figure 3.6 shows both WT and G95R PLK4 expressed within the cells co-localising with both pericentrin and γ-tubulin. The decondensed DNA and the presence of two PLK4 and pericentrin or γ-tubulin puncta within each individual cell confirm that PLK4 is expressed in S phase cells, which is consistent with the onset of kinase activity to initiate centriole duplication (Habedanck *et al.*, 2005). PLK4 expression in G2/M phase cells has also been reported, with an active form of the kinase expressed at M phase, with an assumed role in procentriole assembly and elongation (Sillibourne *et al.*, 2010). However, expression of WT or G95R PLK4 in cells at later stages of the cell cycle was not observed in this study. Future experiments could test the presence of PLK4 at later stages of the cell cycle by arresting the cells in G2/M and performing immunofluorescence with a phospho specific antibody directed toward pSer305, linked to kinase activity. However, these preliminary results confirm that the controlled over-expression of PLK4 using the T-Rex system results in stable expression of exogenous PLK4, which demonstrates similar subcellular localisation at the centrioles as endogenous PLK4.

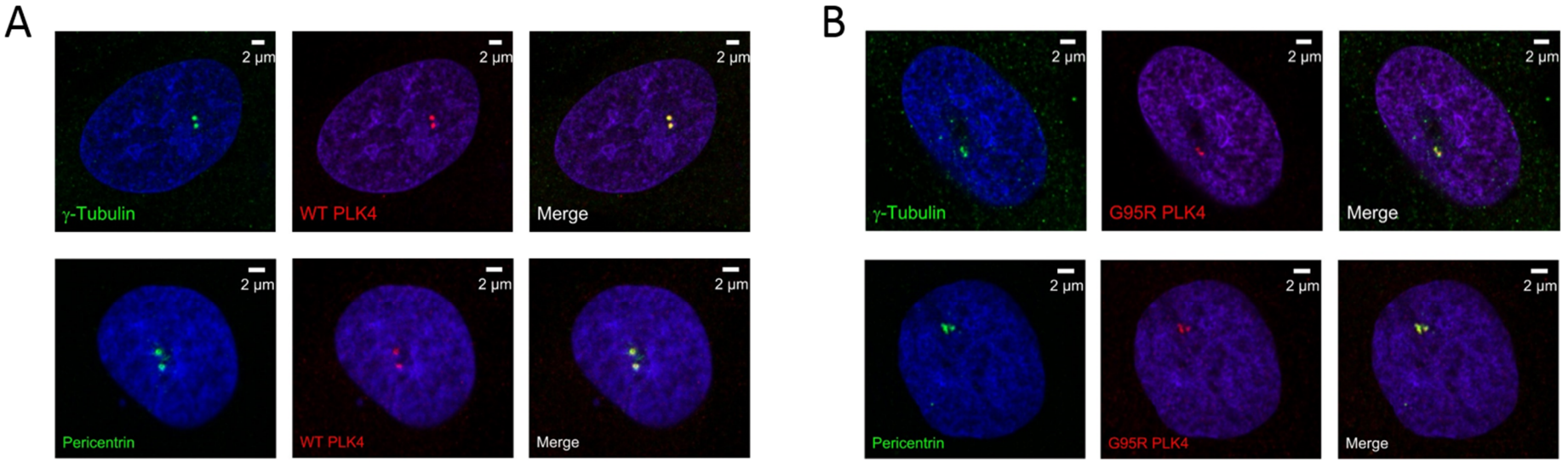


Figure 3.6. WT & G95R PLK4 localise at the centrosome. U2OS cells, induced with tetracycline to express A) myc-WT PLK and B) myc-G95R PLK4, were probed with anti-PLK4 (red) & anti-γ-tubulin or anti-pericentrin (green) antibodies and analysed by immunofluorescence microscopy to confirm co-localisation at the centrosome.

3.2.4. **WT & G95R PLK4 activity assays**

To confidently employ G95R PLK4 as a drug-resistant control to validate PLK4 substrates that are sensitive to inhibition by drugs such as VX-680 and centrinone, the drug-resistant mutant must exhibit similar catalytic activity to the WT kinase. To assess protein activity, recombinant 6His tagged- WT and G95R PLK4 catalytic domains (containing the catalytic domain; residues 1-269) were expressed in Rosetta *E. coli* and purified by nickel affinity chromatography. Analysis of the 12 eluted fractions by SDS-PAGE revealed a high concentration of protein migrating to ~32 kDa, consistent with the size of the catalytic domain of PLK4 (Figure 3.7 A) (purification of G95R PLK4 is shown; WT PLK4 purification is presented in S. Ferries Masters report). Fractions 2-7, containing the highest concentration of protein, were pooled and further purified by gel filtration chromatography on an S200 column (Figure 3.7 B). Fractions were collected across the eluting peak, and re-analysed by SDS-PAGE, which revealed protein of high purity (Figure 3.7 C). The identity of the purified protein was confirmed by digestion with trypsin and analysis by LC-MS/MS (data not shown). Fractions containing the highest concentration of protein were pooled, and characterised using a new *in vitro* PLK4 activity assay.

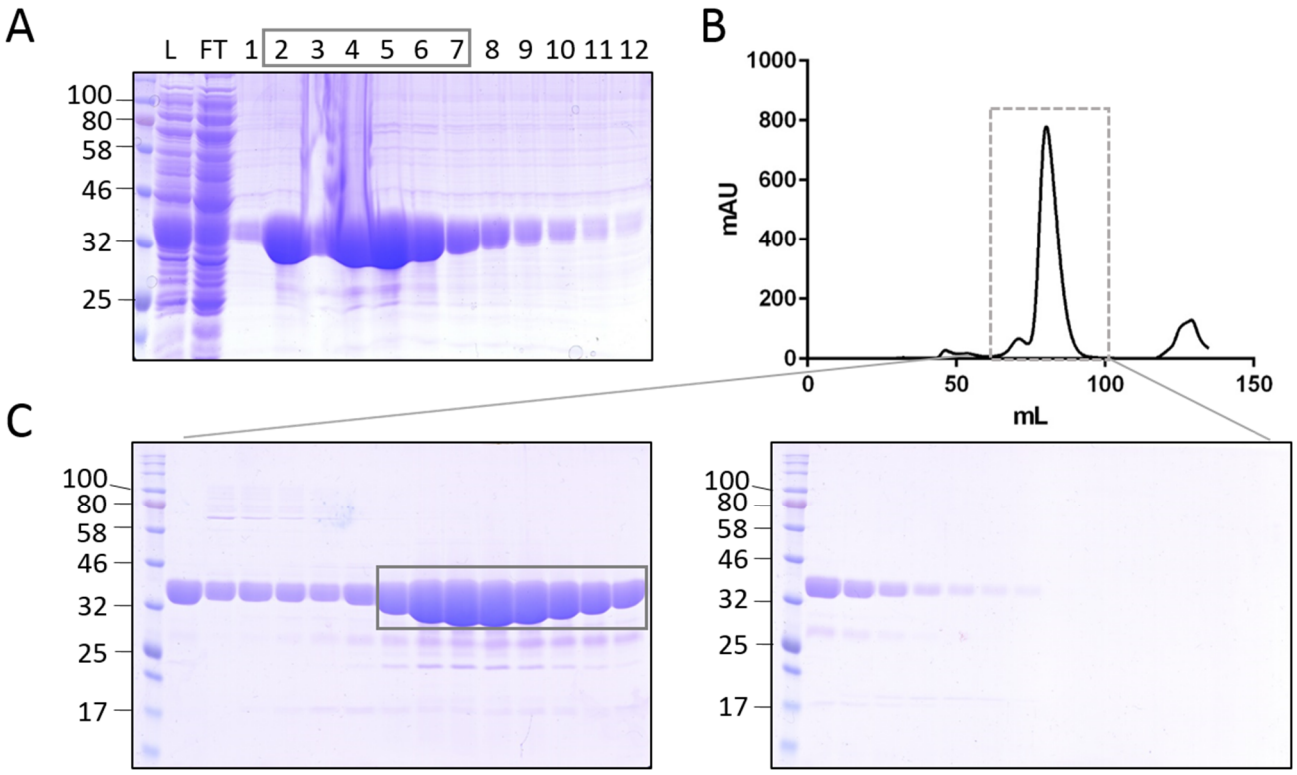


Figure 3.7. Purification of 6His-G95R PLK4 (1-269). A) Following expression in Rosetta *E.coli*, cell lysate was incubated with nickel resin and bound protein was eluted with 500 mM imidazole. 12 fractions were collected and analysed by SDS-PAGE. B) Fractions 2-7 (as shown by rectangle in panel A) were pooled and purified by gel filtration chromatography using a Superdex 200 column. C) Fractions were collected across the main peak and analysed by SDS-PAGE and stained with Coomassie reagent. The rectangle indicates which fractions were pooled for use in *in vitro* assays.

Prior to enzymatic assay, the PLK4 proteins were characterised by a thermal DSF assay to analyse unfolding by incubation with a fluorescent dye capable of binding to hydrophobic regions which become more accessible as the protein unfolds (Rudolf *et al.*, 2014). In addition, binding of inhibitors and the essential co-factors ATP and $Mg^{2+}$was also assessed. The sensitivity of both the WT and G95R PLK4 catalytic domains to the kinase inhibitors VX-680 and centrinone was also assessed. Figure 3.8 shows that both proteins are capable of binding both Mg^{2+} and ATP, with a ΔTm of 7 °C for both WT and G95R PLK4. This is due to a conformational change to the active site induced by binding of Mg^{2+} and ATP, resulting in an inward rotation of the αC helix that creates a more thermostable structure (Knighton *et al.*, 1991; Endicott *et al.*, 2012).

Once bound within the active site, ATP-competitive kinase inhibitors interact through hydrogen bonding with key residues within the protein. Hydrogen bonding between PLK4 and centrinone occurs at residues Met91 within the hinge region and Asp154 of the DFG motif (Wong *et al.*, 2015). These contacts also increase the thermostability of the protein, which can be observed by DSF. The data shows that WT PLK4 appears to bind strongly by both VX-680 and centrinone with ΔTm values of 13 °C and 19 °C respectively (Figure 3.8 A). G95R PLK4 showed very little thermal stabilisation with VX-680 and slight thermal stabilisation in the presence of centrinone (ΔTm ~ 6 °C, Figure 3.8 B), in line with marked drug-resistance in G95R PLK4.

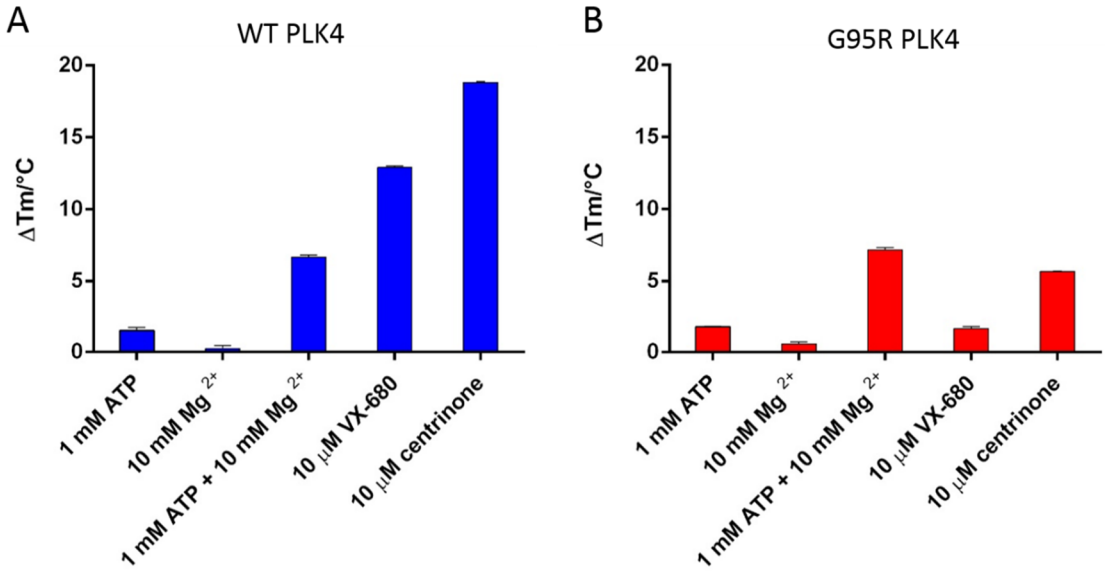
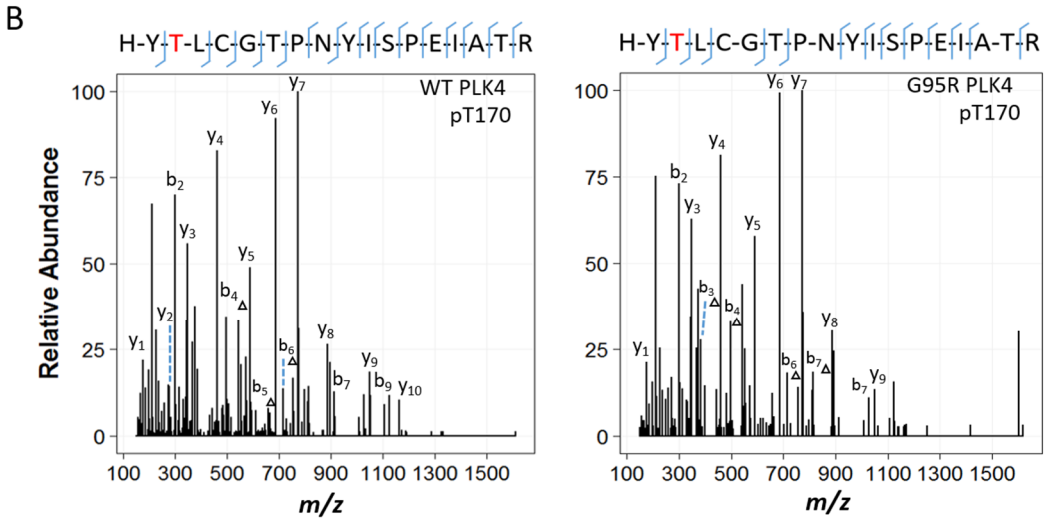


Figure 3.8. Thermal shift analysis of recombinant WT & G95R PLK4 using Differential Scanning Fluorimetry. Thermal shift was measured in the presence of ATP, Mg^{2+} and the potent PLK4 inhibitors, VX-680 and centrinone. The change in Tm (ΔTm) value is reported from three technical replicates of a) WT PLK4 and b) G95R PLK4.

Expression of correctly folded and active protein kinases in *E. coli* often results in extensive autophosphorylation, and PLK4 is known to autophosphorylate *in vitro* (Shrestha *et al.*, 2012). An initial assessment of catalytic activity could therefore be determined for WT & G95R PLK4 by analysing the autophosphorylation status of each purified protein following tryptic digestion. LC-MS/MS analysis of the enriched tryptic phosphopeptides revealed autophosphorylation at 11 sites in WT PLK4, and 10 of the same sites in G95R PLK4, with only phosphorylation of T34 not being observed (Figure 3.9 A). One key autophosphorylation site that was identified was T170, which is located within the activation loop and required for PLK4 activity (Figure 3.9 B) (Swallow *et al.*, 2005; Lopes *et al.*, 2015). In cells, Thr170 autophosphorylation results in activation of PLK4, leading to phosphorylating of the centriolar assembly protein STIL at Ser1108 and Ser1116 within its STAN motif. Phosphorylated STIL then binds SAS-6, initiating centriole duplication (Moyer *et al.*, 2015). Identification of Thr170 autophosphorylation in G95R, with essentially identical autophosphorylation patterns as WT PLK4, strongly suggests that G95R retains WT activity and can therefore be used for the phosphorylation and validation of PLK4 substrates in phosphoproteomics experiments.

A

Peptide Sequence	Site of modification	*m/z*	Charge	Mascot Score	ptmRS score	In WT & G95R?
MATCIGEKIEDFK	T3	811.36	2+	92	100	Yes
GSFAGVYR	S22	468.70	2+	30	100	Yes
AESIHTGLEVAIK	T34	724.36	2+	58	100	Only in WT
DLTLSNLLLTR	T138	669.86	2+	59	100	Yes
IADFGLATQLK	T159	628.82	2+	59	100	Yes
HYTLCGTPNYISPEIATR	T170; Y177; T184	1166.97	2+	77	100; 100; 100	Yes
VVLADYEMPSFLSIEAK	S232	996.48	2+	94	99.8	Yes
LSLSSVLDHPFMSR	S257; S266	874.88	2+	93	100; 100	Yes



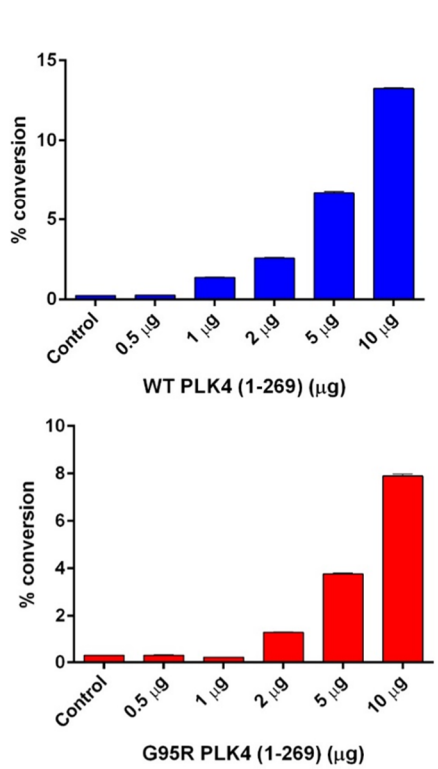


Figure 3.9. Drug-resistant G95R PLK4 retains autophosphorylation activity when expressed in bacteria. A) Full list of autophosphorylation sites identified in WT and G95R PLK4 are depicted with associated *m/z* values, Mascot & ptmRS scores. B) Purified recombinant WT or G95R PLK4 were digested with trypsin and analysed by LC-MS/MS. MS2 spectra generated by HCD shows autophosphorylation at T170 within the activation loop of both WT & G95R PLK4. C) Purified recombinant WT or G95R human PLK4 (amino acids 1–269) was assayed by quantified mobility shift using EZ Reader technology with a PLK4 fluorescent peptide substrate (5-FAM-FLAKSFGSPNRAYKK) in the presence of 1 mM ATP (final concentration), to mimic cellular ATP levels.

An independent assessment of G95R PLK4 activity after purification and storage was undertaken using a kinetic, non-radioactive *in vitro* kinase assay. WT and G95R PLK4 were first incubated with a fluorescent peptide substrate and 1 mM ATP and Mg^{2+} ions, to mimic cellular levels of these co-factors. Conversion of the substrate (5-FAM-FLAKSFG**S**PNRAYKK) into the phosphorylated product (5-FAM-FLAKSFG**S**PNRAYKK) was quantified using a mobility shift assay on the EZ Reader platform. These activity assays reveal that both the WT and G95R protein are active towards a generic PLK4 peptide substrate, with a maximum of 13 % conversion observed for WT and 8 % for G95R over the same time-period (Figure 3.9 C). The very small difference in activity between the two proteins could be due to subtle differences in the tertiary structure arising from the mutation subtly altering the active site.

Further characterisation of the drug-sensitivity of G95R PLK4 was performed by assessing activity toward the peptide substrate in the presence of centrinone, to evaluate the extent of G95R PLK4 drug-resistance. WT & G95R PLK4 were incubated with increasing concentrations of centrinone (10 nM to 100 µM). Figure 3.10 reveals complete inhibition of WT-PLK4 activity at 1 µM compound and a calculated IC_{50} of 330 nM. In comparison, only a small decrease in G95R PLK4 activity is observed, with a maximum decrease of 15 % observed with 100 µM inhibitor. This confirms unequivocally that at high concentrations of ATP, catalytically active G95R PLK4 is highly drug-resistant when compared to WT PLK4, as expected.

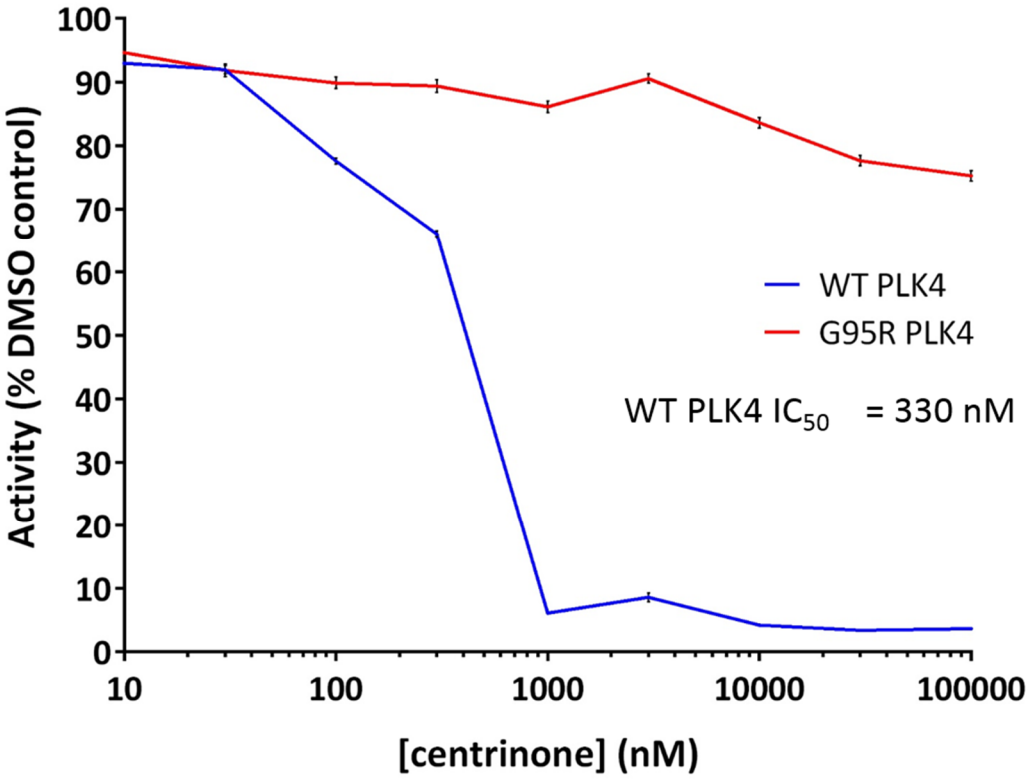


Figure 3. 10. G95R PLK4 is highly resistant to inhibition with centrinone. Purified recombinant WT (blue line) & G95R (red line) PLK4 (1-269) were incubated with fluorescent peptide substrate in the presence of DMSO (control) or the indicated concentrations of centrinone and 1 mM ATP. Reactions were analysed by mobility shift assay on the EZ Reader platform.

An additional assay was performed to test Aurora A activity in the presence of centrinone. Recombinantly produced WT and a drug resistant mutant (G216L) Aurora A (10 ng of each) were assayed against a fluorescent peptide substrate (5-FAM-LRRA**S**LG) in the presence and absence of increasing concentrations of centrinone. Figure 3.11 A shows that Aurora A is sensitive to centrinone, with near complete inhibition of Aurora A activity with 1 µM centrinone. In comparison, a drug resistant Aurora A in which Gly216 has been mutated to Leu shows no sensitivity to centrinone at any of the concentrations tested (Figure 3.11 B). This data further highlights the requirement of a drug resistant PLK4 mutant for validation of PLK4 substrates in a phosphoproteomics study.

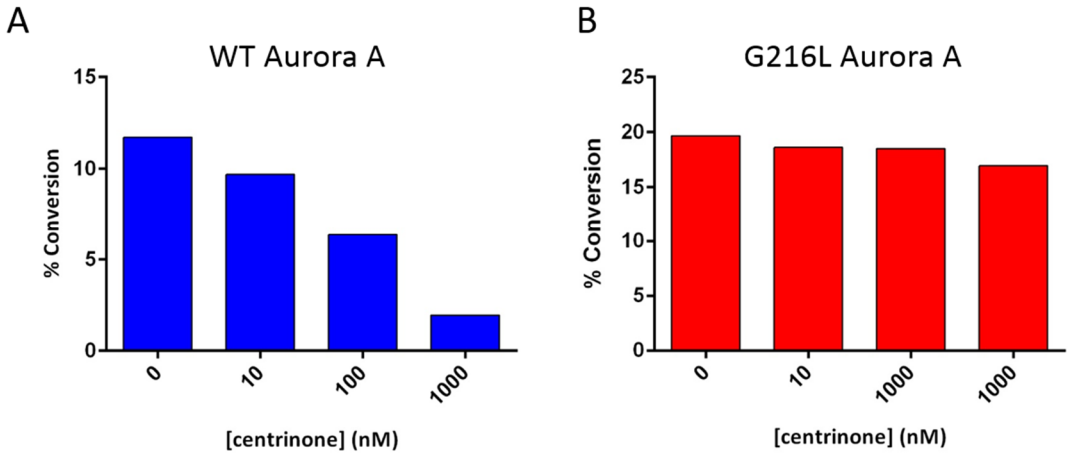


Figure 3.11. Aurora A is inhibited by centrinone. Recombinant WT (A) or G216L (B) human Aurora A was assayed by quantified mobility shift using EZ Reader technology with a fluorescent peptide substrate (5-FAM-LRRA**S**LG) in the presence of 1 mM ATP (final concentration), to mimic cellular ATP levels. Data shown is an average of two replicates.

Taken together, the near identical autophosphorylation sites of WT & G95R PLK4 and the activity of G95R PLK4 against the peptide substrate, even in the presence of high concentrations of centrinone confirm that G95R is a suitable tool to be used for validation of on- and off-target effects of centrinone inhibition of PLK4.

3.2.5. **Cellular PLK4 assays**

Following the development of a suitable cell system for controlled over-expression of PLK4 and characterisation of WT & G95R proteins, the effects of *in vivo* inhibition with centrinone was assessed. Owing to the known promiscuity of kinase inhibitors (Fabian *et al.*, 2005; Bain *et al.*, 2007), incubation conditions were optimised so that the cells could be incubated for the shortest possible time at an inhibitor concentration which was high enough to maximally inhibit PLK4 activity, but that minimises off-target effects.

Initially, U2OS cells expressing WT FLAG-PLK4 were incubated with 150 nM (final concentration) centrinone over a time-course from 1 to 4 hours (Figure 3.12). PLK4 inhibition is reported by its accumulation within the cells, due to the inability of inhibited PLK4 to autophosphorylate its own degron and thus be degraded by the proteasome. The time-dependent inhibition experiment revealed a level of accumulation similar to the MG-

132 control after 4-hour compound exposure. This 4-hour incubation time was therefore selected for the following experiments, since it satisfied the requirement of exposing the cells to centrinone for as short a time as possible to try to limit accumulated off-target inhibitory effects.

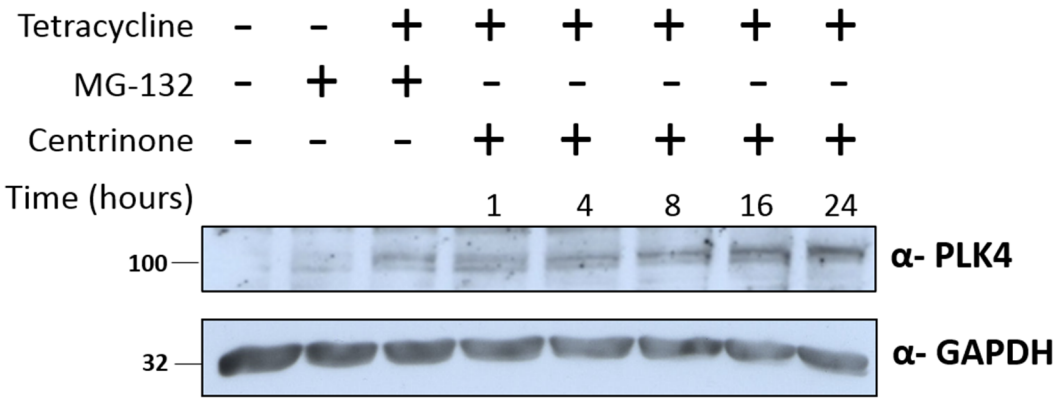


Figure 3.12. FLAG-WT PLK4 is stabilised by centrinone in a time-dependent manner. Expression of FLAG-WT PLK4 was induced upon incubation with 1 µg/mL tetracycline for 18 hours. Cells were incubated with 150 nM centrinone or 10 µM MG-132 for the time points indicated. Total cell lysates were analysed by western blotting using the indicated antibodies.

Next, WT PLK4 expressing cells were incubated with increasing concentrations of centrinone (150 nM to 1,000 nM) for 4 hours. When compared with the levels of PLK4 detected upon treatment with MG-132, an increased accumulation of PLK4 was observed with 150 nM centrinone and stabilisation was essentially maximal at 300 nM (Figure 3.13). No additional accumulation of PLK4 was observed at the higher concentrations of 500 nM and 1,000 nM, when off-target effects are more likely. To confirm these results, total PLK4 was immunoprecipitated from cells and analysed by western blot, which also revealed PLK4 accumulation at 150 nM and 300 nM, but with higher levels of PLK4 observed at 150 nM. Despite this difference, 300 nM was chosen as the optimal concentration to ensure all of the expressed PLK4 was maximally inhibited for phosphoproteomic studies.

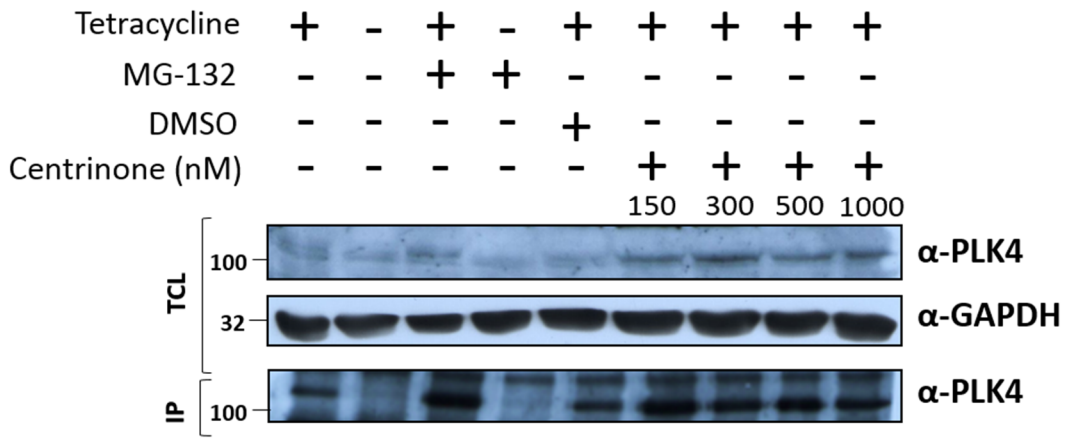


Figure 3.13. FLAG-WT PLK4 is inhibited by centrinone in a concentration-dependent manner. Expression of FLAG-WT PLK4 was induced upon incubation with 1 µg/mL tetracycline for 18 hours. Cells were incubated with centrinone at the concentrations indicated, or 10 µM MG-132, for 4 hours. PLK4 was immunoprecipitated by virtue of its FLAG-tag. Total cell lysates (TCL) were analysed by western blotting using the indicated antibodies

Using the conditions established in the WT PLK4 expressing cells, the experiment was repeated using G95R PLK4 expressing cells to evaluate whether centrinone was able to stabilize the G95R 'drug resistant' protein in cells. Figure 3.14 shows the level of G95R PLK4 accumulation at the drug conditions tested. A slight increase is observed compared to the DMSO control at 150 nM and 300 nM. However, no increase was observed at the higher concentrations of 500 and 1000 nM compared to the DMSO control. However, a decrease compared to the 150 nM centrinone treated cells was observed This suggested that the slight differences were possibly due to uneven transfer of proteins on to the nitrocellulose membrane and not that G95R PLK4 was being inhibited. To confirm this, total G95R PLK4 was immunoprecipitated from cells and analysed by western blot with an anti-PLK4 antibody, which demonstrated that centrinone did not stabilise G95R PLK4, since no accumulation of the protein was observed. Taken together, the time and concentration-dependent experiments established that 300 nM centrinone incubation over 4-hours were suitable conditions to enable a comprehensive analysis of the PLK4 regulated phosphoproteome.

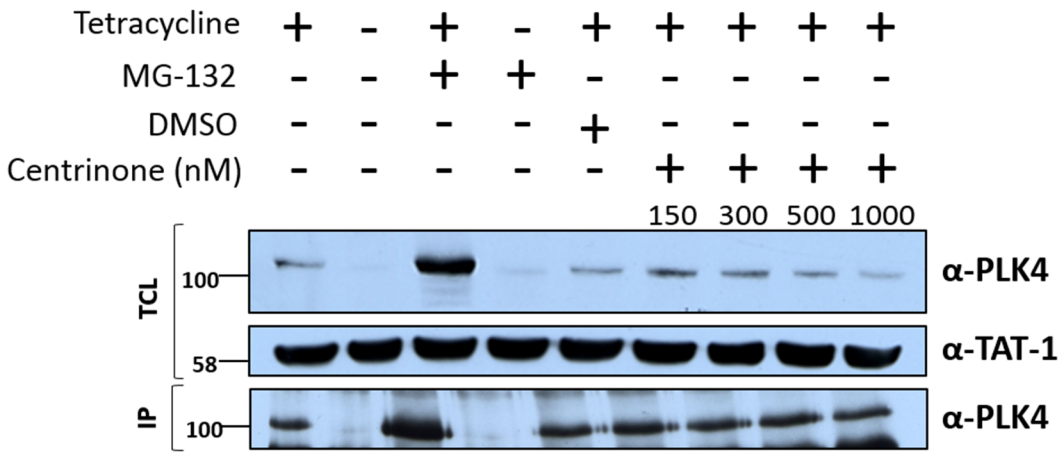
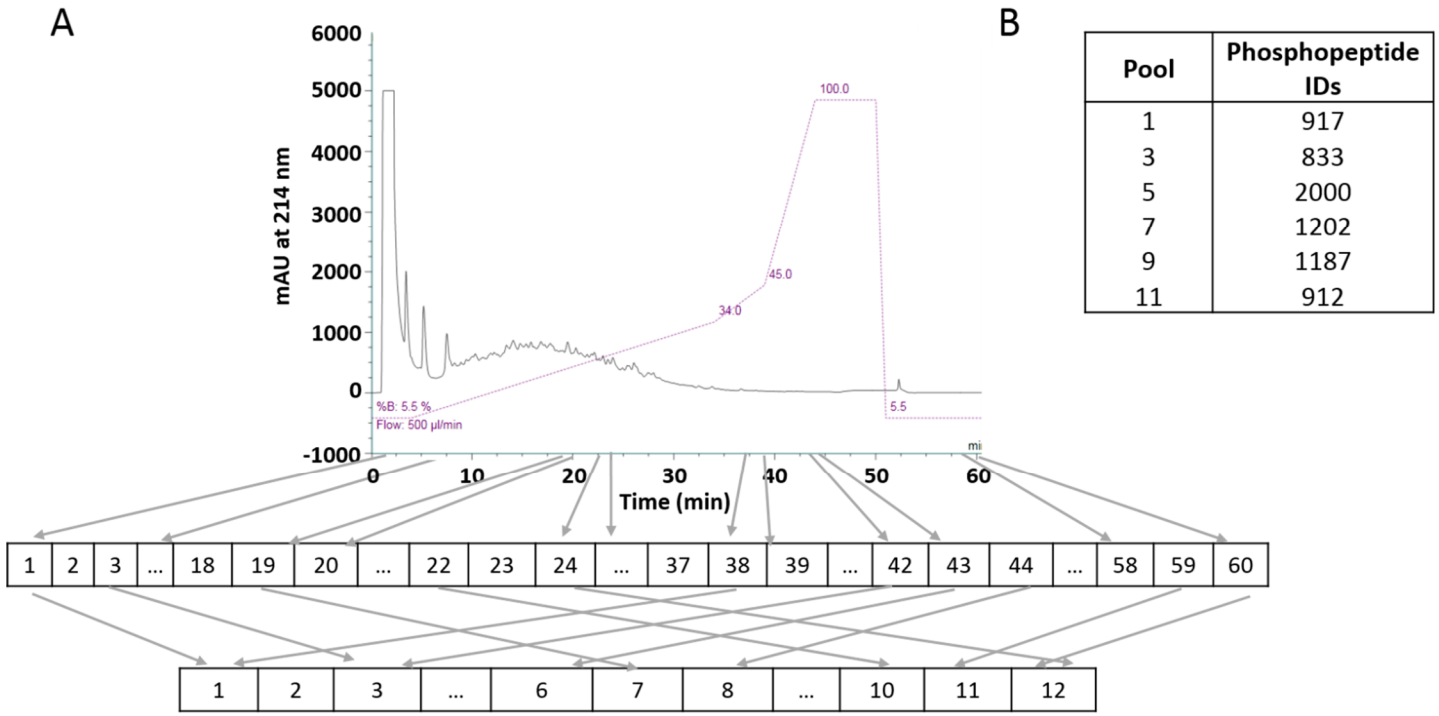


Figure 3.14. FLAG-G95R PLK4 is resistant to centrinone inhibition *in vivo*. Expression of FLAG-G95R PLK4 was induced upon incubation with 1 µg/mL tetracycline for 18 hours. Cells were incubated with centrinone at the concentrations indicated, or 10 µM MG-132 for 4 hours. PLK4 was immunoprecipitated with an anti-FLAG antibody overnight at 4 °C. Total cell lysates and immunoprecipitates were analysed by western blotting using the indicated antibodies.

3.2.6. **High pH reversed phase fractionation**

As discussed in Chapter 1, the study of low abundance proteins and phosphopeptides by LC-MS/MS requires pre-fractionation of the digested peptides as implemented here using high pH reversed phase fractionation, to maximise coverage of the phosphoproteome (Batth *et al.*, 2014) in combination with phosphopeptide enrichment.

As high pH fractionation is performed using a column chemistry that is highly similar to the final chromatographic separation step (low pH reversed phase), a concatenation approach for fraction pooling is necessary (Figure 3.15 A). This approach combines fractions collected from the beginning, middle and end of the high pH separation resulting in pools which, when further separated under low pH reversed phase conditions, most efficiently exploit the separation space of this second dimension, maximising the potential number of identifications that can be made. From the 2.5 mg loaded, 60 fractions were collected across the gradient and concatenated to 12 pools, as shown in Figure 3.15 A. The peptide pools were enriched for phosphopeptides bytitanium dioxide enrichment and analysis of the pools resulted in the identification of similar numbers of phosphopeptides in each, reflecting the success of a concatenation approach to evenly distribute peptides across the gradient for optimal downstream LC-MS/MS phosphopeptide identification (Figure 3.15 B).



Pool	Phosphopeptide IDs
1	917
3	833
5	2000
7	1202
9	1187
11	912

Figure 3.15. Strategy for high pH reversed phase fractionation. A) 2.5 mg of trypsin digested cell lysate was fractionated by high pH reversed phase fractionation. 60 fractions were collected across the gradient and pooled using a concatenation approach to create 12 final pools for subsequent TiO_2 enrichment of phosphopeptides. B) Selected TiO_2 enriched phosphopeptide were analysed by LC-MS/MS and processed in Proteome Discoverer using Mascot with the ptmRS node enabled. The number of phosphopeptides identified in each of the pools analysed (odd number pools) are shown.

3.2.7. **Assessment of SILAC labelling**

Full incorporation of heavy labelled amino acids in SILAC medium is a critical step prior to downstream proteomic analysis. Mann *et al*., suggest that most cell lines will fully incorporate the label after 5 cell doublings and it is critical to confirm full incorporation of the heavy labelled amino acids in to the cells prior to downstream proteomic analysis. WT and G95R FLAG-PLK4 cells were grown in light (R0K0) media and at 80 % confluence were split in to the heavy labelled arginine ($^{13}C_6$ $^{15}N_4$) and lysine ($^{13}C_6$ $^{15}N_2$) media (R10K8). As the cells propagated in the heavy media, aliquots were taken at each cell passage and the cells were lysed, digested with trypsin and analysed by LC-MS/MS to assess incorporation of the heavy label at each passage (Figure 3.16).

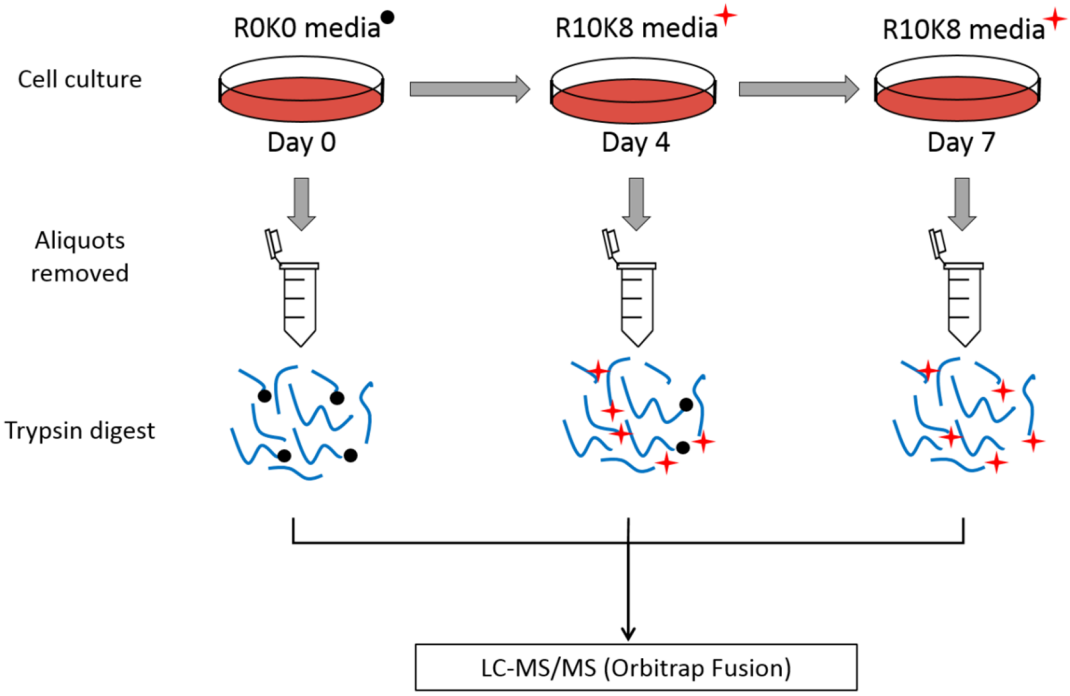


Figure 3.16. Strategy for the assessment of heavy arginine ($^{13}C_6$ $^{15}N_4$) and lysine ($^{13}C_6$ $^{15}N_2$) incorporation in to U2OS cells. Cells grown in 'light' media were subcultured into 'heavy' media containing the stable isotopes arginine and lysine (R10K8). Cell growth and protein turnover results in incorporation of the label in to newly synthesised proteins. At each passage, an aliquot of cells were removed. Proteins were digested with trypsin to produce lysine and arginine terminating peptides and analysed by LC-MS/MS to assess incorporation of the heavy labelled amino acids.

Chromatograms were extracted for specific peptides to assess label incorporation. 95 % incorporation is required for accurate quantification and considered fully labelled, as 100 % cannot be achieved due to impurities in the SILAC medium. Figure 3.17 shows that following approximately 4 cell doublings, the heavy arginine label is partially incorporated, with the light peptide still being observed (20 %) and identified by MS/MS. This incorporation is somewhat slower than suggested by Mann *et al*., and highlights the importance of assessing the rate of incorporation over time for each individual cell line prior to undertaking a large-scale proteomics study. After 7 cell doublings, the cells reached complete incorporation of the heavy labels. A signal for the light peptide (3.6 % of heavy peak) can be observed which is within the accepted range to be considered fully labelled. An assessment of lysine terminating peptides was also analysed to confirm complete incorporation of $^{13}C_6$ $^{15}N_2$ lysine (see Appendix 1). This work established the number of passages required for full incorporation of the heavy labels in U2OS Flp-In cells, and this protocol is therefore suitable for accurate SILAC-based quantitative PLK4 phosphoproteomics experiments in this system.

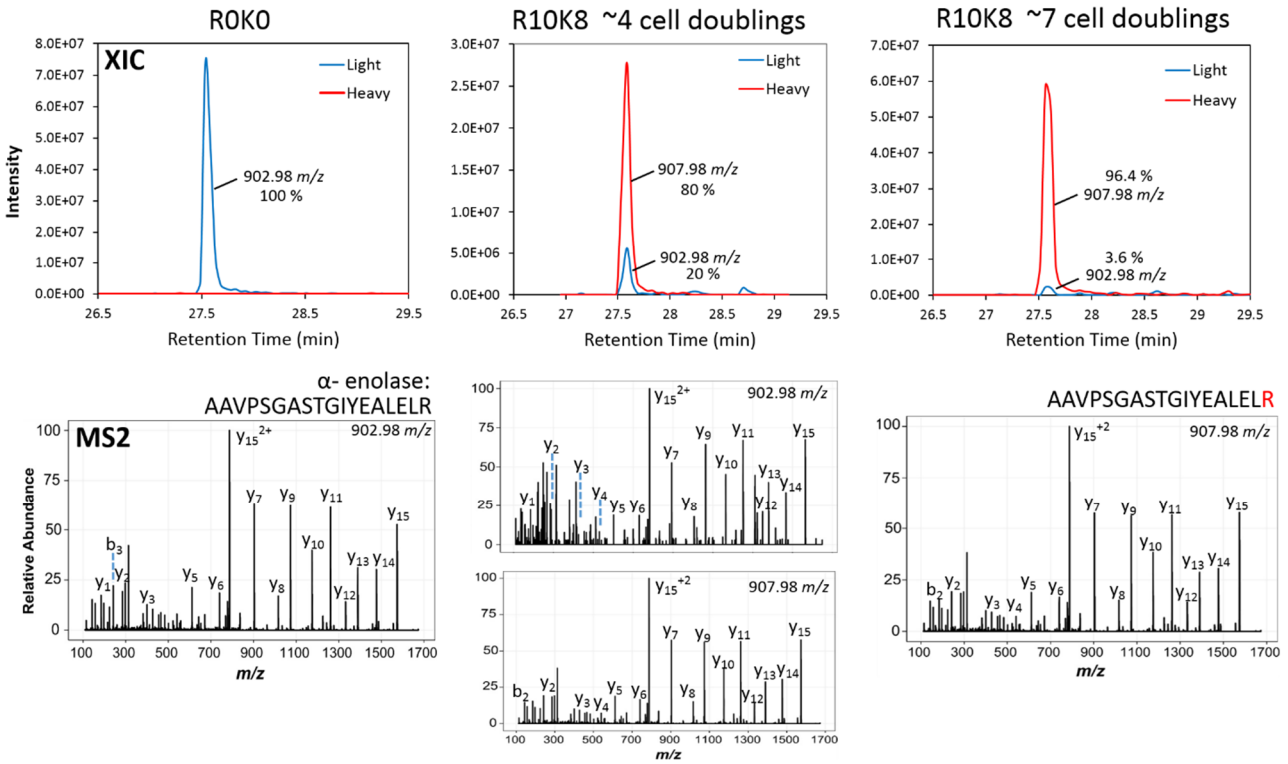


Figure 3.17. Heavy arginine ($^{13}C_6$ $^{15}N_4$) incorporation in U2OS cells. The extracted ion chromatograms (XIC; top panels) show the ion signals for the indicated peptide. The top left panel shows a peptide at *m/z* 902.98 and represents the presence of the peptide in its light form due to growth in 'light' media. MS/MS (bottom left) reveals that the peptide is from α-enolase. Following 4 cell doublings in SILAC (R10K8) media, the same peptide is shown as a SILAC pair, with a mass difference of 10 Da. The 'heavy' peptide is at a higher intensity but still contains sufficient 'light' peptide to be identified by MS/MS (bottom, middle panels), reflecting partial incorporation of the label. At 7 cell doublings, no 'light' peptide can be identified reflecting full incorporation of the heavy label.

3.3. Conclusions

This chapter describes the successful development and optimization of biochemical tools to establish a valid strategy for performing large scale phosphoproteomics to study PLK4 regulated signalling pathways. PLK4 plasmids were successfully cloned into a pcDNA5 vector and subsequently transfected in to U2OS cells to produce stable, isogenic cell lines. These cell lines enabled the controlled over-expression of PLK4, which was not only catalytically active, but was also localised correctly at the centrioles. The generation of these cell lines provides a method for performing a large-scale analysis of PLK4 signalling which has not previously been feasible due to the low abundance of endogenous PLK4 and will permit a true insight in to the activities of PLK4 in the cell.

A key consideration for the work described in this chapter was to establish a viable strategy for validation of phosphoproteomics data obtained for WT PLK4 expressing cells. Probing the cell line with the kinase inhibitor centrinone will perturbate PLK4 regulated signalling pathways but is also likely to cause many off target effects, which could be falsely attributed to PLK4 activity. Confirmation that G95R PLK4 retains WT activity through *in vitro* kinase assays against a peptide substrate and assessment of sites of autophosphorylation was an important result to show that the use of a drug-resistant mutant to validate on and off-target effects of centrinone inhibition of WT PLK4 is a viable strategy.

Finally, optimisation of some of the key stages of the phosphoproteomics workflow, including offline high pH reversed phase fractionation, ensured that during PLK4 phosphoproteomics experiments, relatively high numbers of phosphopeptides could potentially be identified for deep coverage of the phosphoproteome. This allows for the possibility of identifying low abundance proteins that are either direct substrates of PLK4 or implicated in PLK4 regulated signalling. Furthermore, the assessment of SILAC labelling of U2OS cells has revealed the number of cell doublings required for full incorporation of the heavy labels using this cell line which is critical for ensuring accurate and reliable quantification of identified peptides to determine true PLK4 regulated phosphosites and proteins.

The basic tools to enable the study of the PLK4-regulated phosphoproteome by SILAC based quantitative phosphoproteomics are therefore available for the first time. By combining

with a drug-resistance approach, we hoped to use chemical genetics to evaluate PLK4 substrates that were 'on-target' to PLK4. For the most effective exploitation, these tools can be aligned with optimised mass spectrometry methods in order to not only maximise the number of phosphosites identified, but also to have a high degree of confidence in the site of modification. The development and optimisation of this mass spectrometry method is the focus of Chapter 4.

4. Evaluation of parameters for confident phosphorylation site localisation using an Orbitrap Fusion Tribrid mass spectrometer

The text and figures in section 4.2.4 to 4.2.7 are reproduced from the paper 'Evaluation of Parameters for Confident Phosphorylation Site Localization using an Orbitrap Fusion Tribrid Mass Spectrometer' published in Journal of Proteome Research by Ferries *et al*., 2017. I am the first author of the paper and a full transcript, along with supplementary material can be found in Appendix 4. In addition, the Introduction from the paper has been adapted for section 4.1.

4.1. Introduction

Following the development of tools for the functional analysis of PLK4, the next aspect of the project involved optimisation of MS-based methods to maximize coverage of the PLK4 regulated phosphoproteome and achieve high confidence in the sites of phosphorylation. This chapter describes the development and optimization of MS acquisition methods for phosphoproteomics using an Orbitrap Fusion, with a primary focus on the use of mass analysers, fragmentation modes and downstream data processing for phosphopeptide analysis.

Over the past decade, there have been great advances in the analysis of phosphoproteomes by MS, due to the development of highly sensitive mass spectrometers, improved phosphopeptide enrichment techniques and bioinformatics tools for data interrogation. Thousands of phosphopeptides can now be identified in a single analytical 'run'. However, a key challenge that remains to be resolved is the ability to unequivocally identify the site of phosphorylation on a given peptide. Phosphorylation at different sites on a protein can have dramatically different downstream effects, from increasing protein activity, to altering substrate specificity or signalling for protein degradation (Manning *et al.*, 2002). Correct interpretation of the cell signalling pathways under investigation cannot therefore be obtained if there is ambiguity as to the site of phosphorylation within a peptide.

The capabilities of the Orbitrap Fusion Tribrid mass spectrometer, (described in detail in Chapter 1), including the combination of ion trap and orbitrap mass analysers in addition to multiple modes of fragmentation (CID, HCD & ETD) offers significant user flexibility to design methods for the analysis of phosphopeptides. To ensure that the Orbitrap Fusion is utilized to its maximum capability for this type of analysis, key parameters must first be optimized: the HCD collision energy for phosphopeptide fragmentation; maximum fill time for ion accumulation in the orbitrap; fragmentation mode and detection of product ions in the ion trap *versus* the orbitrap. The applied collision energy, along with the peptide sequence and charge, determines the product ions produced following HCD fragmentation. Increased collision energy provides the initial product ions with higher translational energy which results in a greater number of abundant product ions being produced from multiple fragmentation pathways (Olsen *et al.*, 2007). This is crucial for phosphopeptides, where high numbers of product ions are required, not just to identify the peptide but to isolate the site of modification. Collision energy must also be controlled to ensure over-fragmentation does not occur; higher energies can result in multiple fragmentation events, decreasing the abundance of the larger *b*- and *y*- ion fragments, resulting in spectra that are dominated by smaller *m/z* ions and internal product ions. This can decrease the overall quality of the spectra, resulting in fewer identifications (Diedrich *et al.*, 2013). .

In addition to collision energy, the maximum injection times for ion accumulation in the orbitrap for MS2 must also be optimised for phosphoproteomics studies to achieve a balance between quantity and quality of acquired spectra. Methods are created with automatic gain control (AGC) to fill the orbitrap until a set number of ions are reached and with a time limit imposed for ion accumulation. Owing to the relatively low abundance and reduction in ionization efficiency that can be observed for some phosphopeptides, the AGC is unlikely to be reached in most instances without long maximum injection times. This was shown in a previous study which revealed that the average injection time for phosphopeptides in an orbitrap was 236 ms and as a result, fewer phosphopeptides were identified than when using CID based fragmentation in a linear ion trap (Zhang *et al.*, 2009). Whilst this is dependent on the material loaded, it does reflect the longer injection times required for orbitrap mass analysis. Additionally, a study on the Q Exactive platform, which assessed a range of methods with varying injection times for peptide identifications revealed that whilst ‘fast’ methods using low injection times produced high scan numbers and high quality data for peptide analysis, much longer fill times were required to study

post-translationally modified peptides (Kelstrup *et al.*, 2014). Maximum injection time is therefore an important parameter to ensure sufficient ion accumulation to obtain high quality MS/MS spectra whilst maintaining a high number of phosphopeptide identifications.

Increasing the number of site-determining product ions can be achieved by the exploitation of multiple complementary fragmentation modes (described in Chapter 1) (Boersema *et al.*, 2009a; Kim *et al.*, 2011; Frese *et al.*, 2012; Frese *et al.*, 2013; Lanucara *et al.*, 2014; Lanucara *et al.*, 2016). On the Orbitrap Fusion, the benefit of being able to perform both HCD and CID, as well as ETD and ETca/hcD, with product ion analysis being performed in either the ion trap or the orbitrap (Hebert *et al.*, 2014; Riley *et al.*, 2016; Espadas *et al.*, 2017), means that these instruments should be of great benefit to improve phosphoproteome analysis and unambiguous phosphosite identification.

The number of potential phosphopeptide MS acquisition strategies, particularly with the new generation of versatile tribrid Orbitrap instruments, means that it can be extremely complicated and time-consuming to establish an 'optimal' phosphoproteomics pipeline. There are numerous challenges associated with optimizing instrument settings to maximize phosphopeptide identification and crucially, confident site localization. The added capability of the Orbitrap Fusion instruments to parallelize acquisition of MS1 in the high resolution Orbitrap, while acquiring at a faster rate, lower resolution MS2 in the ion trap (if required), means that there can be significant advantages for high throughput proteomics using this type of tribrid instrument. The number of possible strategies for MS(/MS) data acquisition (orbitrap *versus* ion trap), as well as potential fragmentation regimes (CID, HCD, ETcaD, EThcD, with or without neutral loss considerations that may be used for triggering additional MS2 /MS^3 acquisition, or multistage acquisition (MSA)) means that the combinatorial options for MS data acquisition are vast.

4.1.1. **Aims**

The aims of the work described in this Chapter were therefore, to develop and optimize MS parameters and a downstream data analysis workflow for the study of the PLK4-regulated phosphoproteome. This included preliminary experiments to ascertain a suitable collision energy and MS2 maximum injection time for phosphopeptides, followed by the evaluation of eight acquisition modes on the Orbitrap Fusion Tribrid MS platform, using a library of synthetic phosphopeptide standards, and a complex phosphopeptide-enriched cell lysate

preparation. In addition, this work aimed to identify optimal downstream processing for phosphopeptide identifications and phosphosite localization by comparing Proteome Discoverer (PD) with MASCOT and phosphoRS (*ptmRS*), and MaxQuant with Andromeda and PTM-score.

4.2. Results & Discussion

4.2.1. Effect of normalized collision energy on phosphopeptide identifications

To assess the effect of HCD collision energy on phosphopeptide identification and confident phosphosite localisation, a phosphopeptide enriched U2OS lysate was analysed on the Orbitrap Fusion, with both MS1 & MS2 scans being performed in the orbitrap to generate high-resolution data. A range of collision energies were tested from 28 to 34 normalised collision energy (NCE), and the data processed using Mascot within Proteome Discoverer, with the ptmRS node enabled. Table 4.1 shows the total number of phosphopeptide identifications for each method tested, with the total number of identifications increasing by ~30 % from 28 to 32 NCE. The numbers then decrease at 34 NCE, suggesting that at this collision energy, over fragmentation may be reducing MS2 spectral quality resulting in fewer phosphopeptide identifications.

Table 4.1. Evaluation of normalised collision energy (NCE) on the number of phosphopeptides identified. For each NCE value tested, the number of PSMs, total number of peptide and phosphopeptide identifications, and number of phosphosites is reported. 32 NCE, which identifies the highest number of phosphopeptides is highlighted in red.

NCE	Total PSMs	Total IDs	Phosphopeptide IDs	No. of phosphosites
28	7373	3495	2501	2685
30	10354	4591	3226	3513
32	**10273**	**4569**	**3277**	**3586**
34	10067	4427	3137	3370

4.2.2. **Effect of normalised collision energy on phosphosite confidence**

To assess the effect of collision energy on phosphosite localisation, individual spectra were manually annotated to identify the number of site determining ions produced following fragmentation at each NCE value tested. Figure 4.1 shows annotated MS2 spectra for the peptide HNsESESVPSSMFILEDDR from 28, 30 & 32 NCE. At 28 NCE, the fragment ions identified provide evidence for phosphorylation at either the first, second or third serine in the sequence. However, insufficient fragment ions are produced to unequivocally isolate the phosphosite. This is reflected in the ptmRS score of 33.3 for this peptide. At 30 NCE, the b_6 ion is produced which eliminates the third serine as the site of modification and improves the ptmRS score to 50. However, the exact site of phosphorylation still cannot be determined. The b_4-H_2O product ion observed when 32 NCE is applied isolates the first serine in the sequence as the phosphorylated residue and a ptmRS score of 100 is obtained.

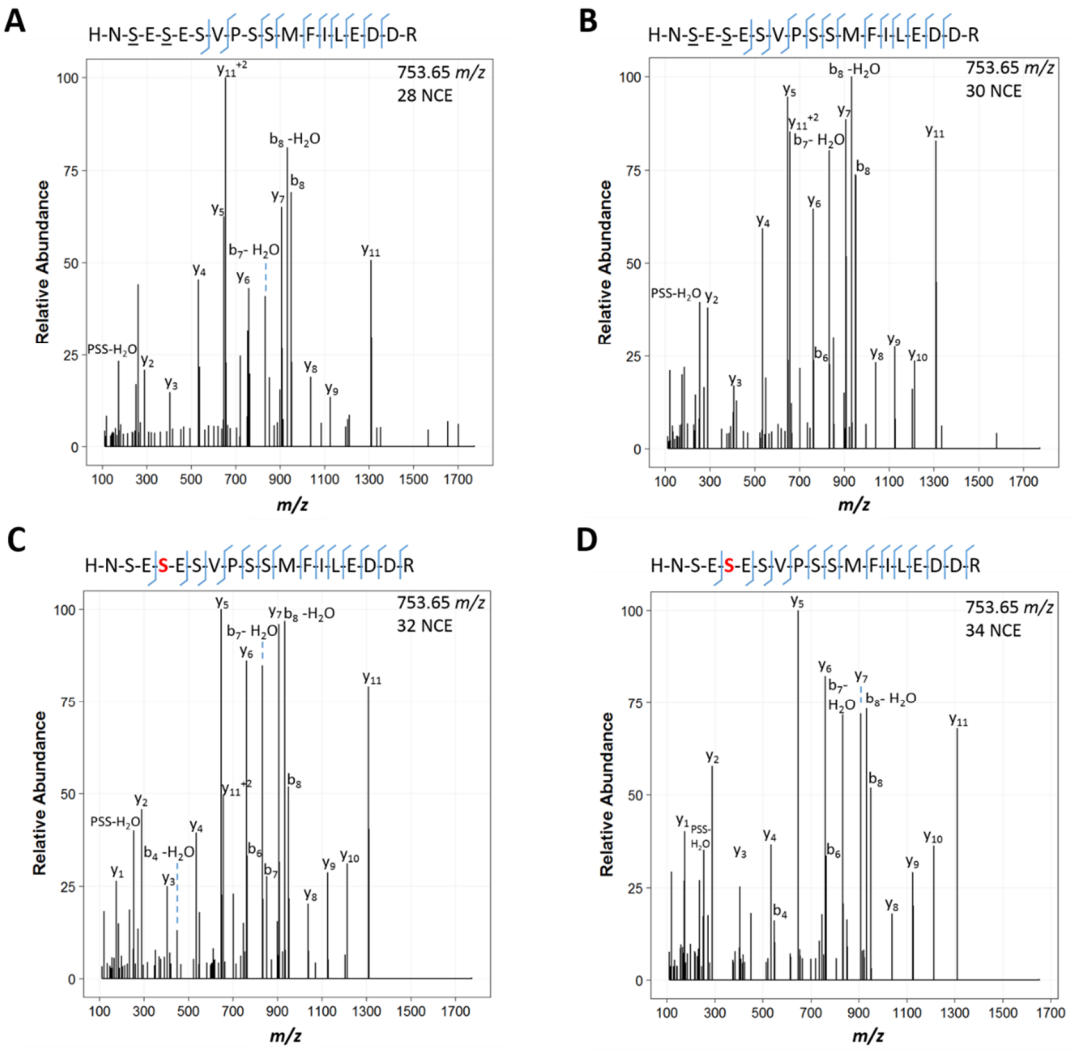


Figure 4.1. Product ion spectra of a phosphopeptide fragmented with 28, 30, 32 & 34 NCE. Doubly charged ion at *m/z* 753.65 fragmented at the indicated collision energies: A) 28 NCE; B) 30 NCE; C) 32 NCE; D) 34 NCE. The sequence is displayed on each mass spectrum showing the product ions identified and isolation of the phosphosite.

Next, the phosphopeptides identified under all four conditions were compared to assess the global distribution of ptmRS scores as NCE increased. Interestingly, the data shows that the percentage of high confidence 'Type I' phosphosites (ptmRS ≥ 75 %) was relatively high across all collision energies tested (~82 %) (Figure 4.2A). Figure 4.2B shows that the distribution of ptmRS scores does not change with increased collision energy and that the vast majority of phosphosites are given a confidence score above ~98 % or below ~50 %. It is likely that the increased number of phosphopeptides identified as collision energy increases is due to accessing a wider population of peptides that could be fragmented to produce sufficient fragment ions for identification, but the ability to localise phosphosites follows the same distribution regardless of the energies tested in this study.

A

Collision Energy	No. of phosphosites	ptmRS ≥0.75	ptmRS ≥0.75 (%)
28	2685	2201	82%
30	3513	2875	82%
32	**3586**	**2972**	**83%**
34	3370	2753	82%

B

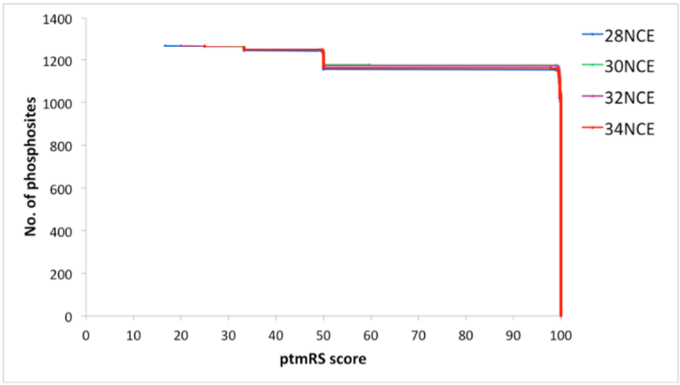


Figure 4.2. Evaluation of collision energy on ptmRS score. A) The total number of phosphosites identified and those identified with high confidence ptmRS (≥0.75, is indicated. 32 NCE, which identifies the highest number of confident phosphosites is highlighted in red. B) Global distribution of ptmRS scores for phosphosites identified in all four collision energy methods tested**.**

4.2.3. **Effect of maximum injection time on phosphopeptide identifications & phosphosite confidence**

Next, an assessment of the effect of maximum injection time of ions into the orbitrap on phosphopeptide identifications and phosphosite localisation confidence was performed. A range of MS2 fill times from 50 ms to 100 ms were tested and the data processed using Mascot/ Proteome Discoverer with phosphosite localisation assessed using ptmRS.

Table 4.3A shows that as the MS2 maximum injection time increases, the number of both MS1 and MS2 scans decreases, as expected. However, the number of phosphopeptides identified increases 10 % as fill time increases from 50 ms to 80 ms (Table 4.3B), demonstrating that despite obtaining fewer spectra, the quality of the acquired spectra is higher and therefore more phosphopeptides can be identified with longer fill times.

Interestingly, whilst a slight increase in the number of singly phosphorylated peptides is observed as maximum fill time increases, a greater difference is observed for multiply phosphorylated peptides (Figure 4.2B). At 100 ms injection time, the number of doubly and triply phosphorylated peptides increased by 25 % and 64 % respectively, compared with the 50 ms injection time method. Whereas reduced ionization efficiency for singly and doubly phosphorylated peptides has been disputed (Steen *et al.*, 2006), multiply phosphorylated peptides have been shown to suffer disproportionally from ionization suppression. Efforts have been made to increase the identification of these multiply phosphorylated peptides, including by addition of EDTA and phosphoric acid to LC buffers to aid elution from reversed phase columns (Kim *et al.*, 2004; Fleitz *et al.*, 2013). The increase in the number of multiply phosphorylated peptides identified with longer injection times without the need for additives therefore presents an ideal approach for a global analysis of the phosphoproteome, including this subset of phosphopeptides that would not otherwise be observed.

Whilst the numbers of identified triply phosphorylated peptides remains low with 100 ms injection time (80 phosphopeptides), this is to be expected, as identification of multiply phosphorylated peptides requires there to be an equal number of basic residues to phosphate groups which does not occur frequently when trypsin is used for digestion. Digestion with another enzyme, such as LysC which would only cleave C-terminal to lysine residues, resulting in a number of basic peptides including both arginine and lysine to

counter the acidic effects of the phosphate on modified residues would allow for potential identification of a greater number of multiply phosphorylated peptides.

When considering ptmRS score, the percentage of confidently localised phosphosites (≥ 75 %) increases from 81 to 84 % for 80 ms, and 86 % for 100 ms fill times. The global distribution of ptmRS scores was assessed and clearly shows that with longer MS2 fill times the ptmRS score increases up to 80 ms (Figure 4.3C). Again, this is likely due to longer fill times permitting the accumulation of sufficient numbers of ions resulting in higher quality MS2 spectra upon fragmentation. Overall, 100 ms injection time represents a significant improvement from the lower injection time methods tested and allows for a high number of quality spectra to be identified and is optimal for a global phosphoproteomics study.

A

MS2 Fill Time	MS1 Scans	MS2 Scans
50	10212	55701
70	9171	48019
80	8707	44851
100	8050	39181

B

MS2 Fill Time	Total IDs	Phosphopeptide IDs	1 Phospho	2 Phospho	3 Phospho	No. of phosphosites	ptmRS ≥0.75	ptmRS ≥0.75 (%)
50	6184	4190	3657	502	29	4744	3855	81%
70	6544	4555	3892	610	51	5262	4427	84%
80	6691	4708	3992	657	55	5457	4599	84%
100	**6435**	**4596**	**3841**	**669**	**80**	**5417**	**4660**	**86%**

C

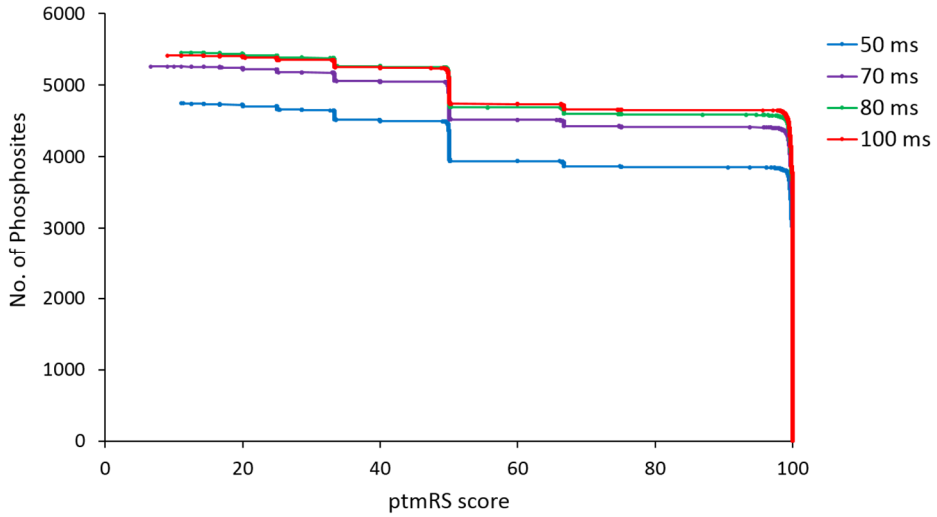


Figure 4.3. Evaluation of MS2 fill time maximum injection time on phosphopeptide identification and phosphosite confidence. A) MS2 fill times tested and the respective number of MS1 and MS2 scans acquired. B) The total number of peptides & phosphopeptides identified is shown in addition to the number of singly, doubly and triply phosphopeptides identified. Number of phosphosites and phosphosite confidence is indicated. 100 ms injection time, which identified the highest number of confident phosphosites is highlighted in red. C) Global distribution of ptmRS scores is shown for each method.

Taken together, the results show that for high numbers of phosphopeptide identifications, including doubly and multiply phosphorylated peptides, and high confidence in the site of modification, a collision energy of 32 NCE and a maximum MS2 fill time of 100 ms provides optimal parameters for large-scale phosphoproteomics analysis.

4.2.4. **Comparison of fragmentation methods and MS2 resolution settings for identification and site localization of phosphopeptide standards**

To evaluate the advanced capabilities of the Orbitrap Fusion tribrid mass spectrometer for site-specific phosphopeptide identification, a series of MS acquisition methods were designed to assess the benefits of using either the high resolution orbitrap or the lower resolution ion trap mass analysers. In the first instance a commercially available synthetic library of phosphopeptides was analysed, (Marx *et al.*, 2013) that comprised tryptic peptides previously observed in multiple large-scale phosphopeptide studies. The library was designed such that the typical composition and length observed in bottom-up proteomics is represented, with a natural occurrence of unmodified and phosphorylated serine, threonine and tyrosine residues.

As well as differing in resolving power, there are significant differences in speed and sensitivity between the orbitrap (OT) and ion trap (IT) mass analysers. HCD, EThCD and neutral loss (NL) triggered ETD-mediated fragmentation strategies, where ions exhibiting precursor neutral loss of 98 amu (arising due to the characteristic loss of H_3PO_4 from phosphorylated peptide ions (DeGnore & Qin, 1998; Boersema *et al.*, 2009a; Lanucara *et al.*, 2014)) or 80 amu (arising due to loss of HPO_3) following HCD, were also compared (Table 4.2; Table S1).

Table 4.2. MS data acquisition methods evaluated. IT: ion trap; OT: orbitrap; nl: neutral loss. See Appendix 4, Table S1 for full details of MS data acquisition parameters. All MS1 analysis was performed in the orbitrap.

Method	Resolution (MS1)	Mass analyzer (MS2)	Resolution (MS2)
HCD OT	60K	Orbitrap	30K
HCD IT	120K	Ion Trap	Rapid
EThcD IT	120K	Ion Trap	Rapid
EThcD OT	60K	Orbitrap	30K
HCD OT nl EThcD IT	60K	Orbitrap	30K
HCD OT nl ETcaD IT	60K	Orbitrap	30K
HCD IT nl EThcD IT	120K	Ion Trap	Rapid
HCD IT nl ETcaD IT	120K	Ion Trap	Rapid

The phosphopeptide library, containing 175 unique phosphopeptides (191 phosphorylation sites), was divided into five pools for LC-MS/MS analysis (see Appendix 2). Isomeric phosphopeptides (where the same peptide sequence is modified on a different residue) were allocated to different analytical pools to ensure that site localization could be defined absolutely. The five pools of synthetic phosphopeptide standards were each analysed in duplicate using the eight MS acquisition methods, assessing both phosphopeptide identification and phosphosite localization (Table 4.2, Appendix 4 , Supp. Fig 1).

As an extension of previously published studies (Marx *et al.*, 2013; Wiese *et al.*, 2014) the ability of two commonly used phosphoproteomics data analysis platforms, MASCOT integrated into Proteome Discoverer (PD) using ptm*RS* (a slightly modified version of phospho*RS* (Taus *et al.*, 2011)) for phosphosite localization, and Andromeda with MaxQuant and PTM-score (Olsen *et al.*, 2006) was also assessed, to identify the synthetic phosphopeptides from all eight datasets (Table 4.3).

Table 4.3. MS acquisition and data analysis methods evaluated using synthetic phosphopeptides. For each of the eight orbitrap Fusion MS acquisition methods (Table 1, Supplementary Table 1) the number of peptide spectrum matches (PSMs) are presented (n = two technical replicates), together with the number of unique peptides (out of a total of 175) and phosphosites (total 191), as well as the number and percentage of correctly localized phosphosites using either Andromeda with PTM-score (top), or MASCOT and ptmRS (bottom), according to the top-ranked PSM. [a]Mean values are presented ± S.D.

Search Engine		HCD OT	HCD IT	EThcD OT	EThcD IT	HCD OT nl EThcD	HCD OT nl ETcaD	HCD IT nl EThcD	HCD IT nl ETcaD
Andromeda	**# PSM[a]**	705 ± 4	984 ± 16	407 ± 18	515 ± 88	625 ± 194	650 ± 30	838 ± 37	745 ± 36
	# unique phosphopeptides	154	159	146	152	156	154	154	155
	# phosphosites	168	173	160	166	170	168	168	170
	# phosphosites correctly localized with PTM-score	155	150	155	156	152	154	147	150
	% phosphosites correctly localized with PTM-score	92%	87%	97%	94%	89%	92%	88%	88%
MASCOT	**# PSM[a]**	889 ± 1	1497 ± 11	417 ±1	654 ± 52	866 ± 107	868 ± 13	1029 ± 30	940 ± 29
	# unique phosphopeptides	164	168	149	156	162	160	159	157
	# phosphosites	179	183	163	165	180	175	173	171
	# phosphosites correctly localized with *ptm*RS	172	151	154	151	167	163	154	144
	% phosphosites correctly localized with *ptm*RS	96%	83%	94%	92%	93%	93%	89%	84%

Implementation of either the Andromeda or MASCOT search algorithms resulted in notably fewer PSMs using EThcD compared to HCD, independent of whether MS2 was performed in the orbitrap or the ion trap (Table 4.3). This result can be explained by the increase in duty cycle for this mixed mode fragmentation regime. Consequently fewer phosphopeptides were identified with EThcD OT compared with the analogous HCD OT, and likewise for EThcD IT compared with HCD IT (Table 4.3). However, the higher percentage of PSMs with correctly localized phosphosites following EThcD IT (94% compared with 87% for Andromeda/PTM-score; 92% compared with 83% for MASCOT/ptm*RS* for EThcD IT or HCD IT respectively) translated to the same or higher numbers of correctly site localized phosphosites being characterized overall with EThcD IT than HCD IT (Table 4.3; Figure 4.4A; Appendix 4 Supp. Figure 2). These findings are in agreement with previous observations on different instrument platforms, which highlight the benefit of mixed mode fragmentation for improved phosphosite localisation (Frese *et al.*, 2013). For the high resolution OT data, there was a notable difference in the performance of the two search engines. Consequently, while phosphosite localization confidence increased with EThcD compared with HCD (resulting in the same numbers of correctly localized phosphosites) using Andromeda/PTM-score, this was not the case with MASCOT/*ptmRS*. 172 phosphosites were correctly identified with HCD OT, whereas EThcD OT yielded only 154 correctly localized phosphosites. The benefits of high resolution MS2 acquisition therefore appear to outweigh the increased duty cycle associated with EThcD when using MASCOT/*ptmRS* for this phosphopeptide library.

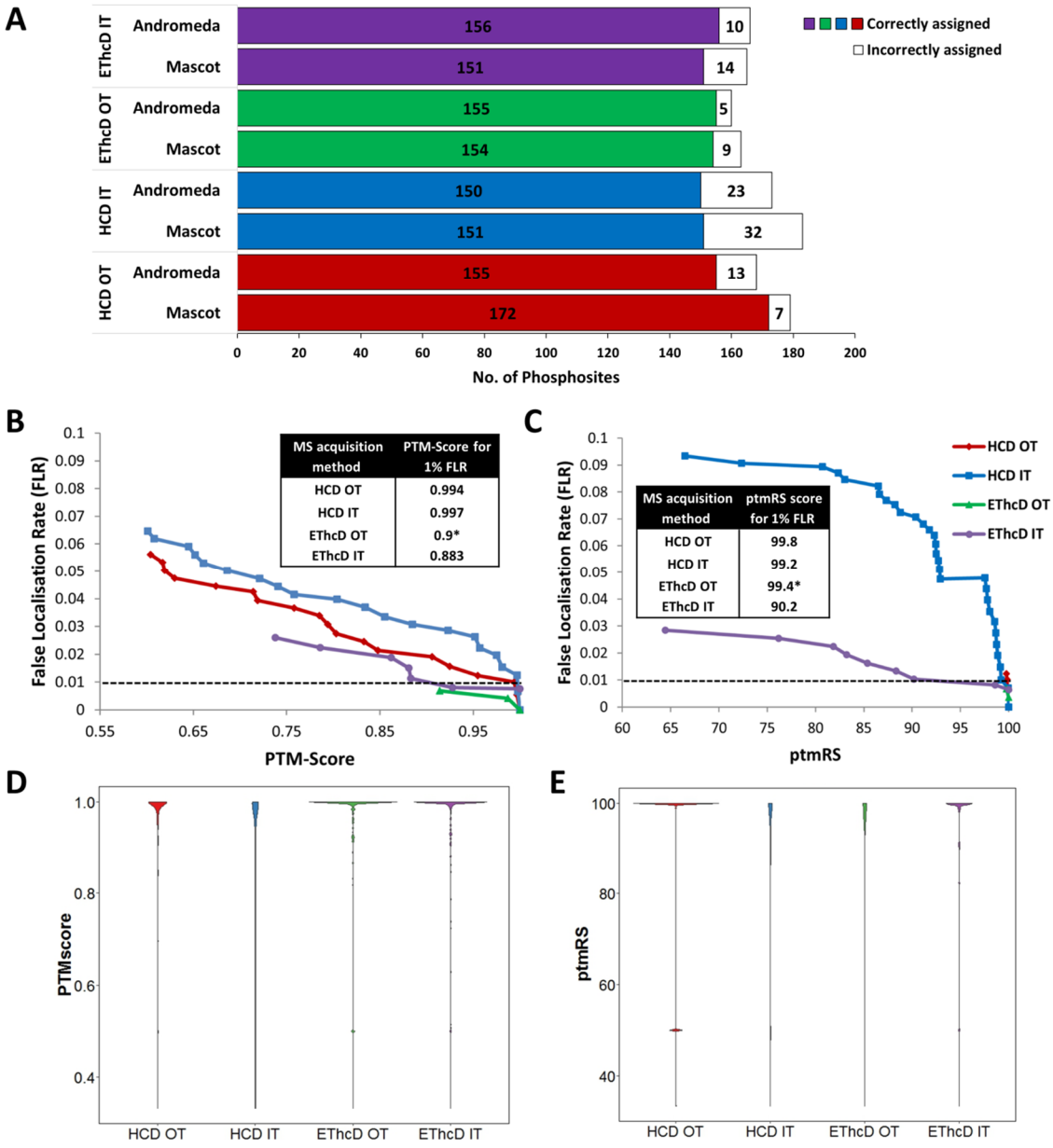


Figure 4.4. Fragmentation method-specific phosphosite localization. (A) Number of correctly assigned (HCD OT: red; HCD IT: blue; EThcD OT: green; EThcD IT: purple) and incorrectly assigned (white) phosphosites from the synthetic phosphopeptide library (Table 1; Supplementary Table 1). (B, C) False localization rate (FLR) determination for the four different basic MS2 acquisition strategies using either Andromeda/PTM-score (B) or MASCOT/ptmRS (C). (D, E) Distributions of PTM-score (D), or ptmRS (E) for each of the four MS2 methods. *Site localization scores equivalent to 0.7% FLR

When considering HCD fragmentation, with or without NL-triggered ET(hc/ca)D, phosphosite localization with both bioinformatics platforms was optimal (higher percentage) with high resolution orbitrap MS2 analysis, likely due to the improved confidence afforded by the enhanced mass accuracy as compared with low resolution ion trap MS2 measurements (Table 4.3; Figure 4.4A; Appendix 4 Supp. Figure 1). Interestingly, Andromeda/PTM-score yielded fewer numbers overall, both of unique phosphopeptides and correctly localized phosphosites, compared with MASCOT/*ptmRS*, irrespective of MS method. A maximum of 159 unique phosphopeptides (155 correctly localized phosphosites) were identified from the pool of 175 synthetic phosphopeptides with Andromeda/PTM-score, compared with 168 phosphopeptides (172 correctly localized phosphosites) when the same data were interrogated using MASCOT/*ptmRS*.

With both search algorithms, HCD IT was optimal for both PSMs and the numbers of unique phosphopeptide identified, as might be expected given the possibility for parallelization of MS1 data acquisition in the orbitrap and concurrent MS2 analysis in the ion trap. However, site localization confidence, the critical parameter from the point of view of biological inference, was either optimal (*ptmRS*) or of equal performance (PTM-score) using the HCD OT method.

Upon further examination of the workflows exploiting neutral loss-triggered ETcaD, the vast majority (89 – 93%) of correctly site localized phosphopeptides were derived from the HCD spectra rather than the ETcaD spectra triggered following precursor neutral loss. The additional incorporation of ETcaD in this regime thus appeared to offer no benefit for either phosphopeptide identification or site localization over that achieved with HCD alone. Indeed, the number of PSMs was compromised due to the increase in duty cycle for the EThcD component of this multi-stage acquisition method. The HCD IT/OT nl ETcaD IT methods are therefore not discussed in subsequent analytical comparisons.

A significant advantage of using synthetic phosphopeptides of known sequence is the ability to define false localization rates (FLRs) specific to the MS acquisition method employed, by counting the numbers of correct and incorrectly site localized PSMs (Marx *et al.*, 2013) (Figure 4.4 B-E). The distribution of site localization scores for each of the four unique fragmentation modes, HCD OT, HCD IT, EThcD OT, EThcD IT, with each of the two informatics pipelines is presented in Figure 4.4. Akin to previous observations on different

MS platforms with both synthetic phosphopeptides (Marx *et al.*, 2013) and a complex phosphopeptide enriched cell lysate (Wiese *et al.*, 2014), both site localization tools require MS acquisition method specific scores to yield a 1% FLR, (Figure 4.4 B, C). It is of interest to note that although fewer phosphosites were incorrectly localized overall with HCD OT compared to HCD IT with both search engines, this does not correlate with a lower site localization score. *ptmRS* exhibits a bimodal distribution for high-resolution MS2 data, with clustering of values around *ptmRS* = 100 and *ptmRS* = 50, indicating either 'certainty', or lack of discriminatory evidence between two possible sites, respectively. In contrast, PTM-score values are more evenly distributed (red plots in Fig. 4.4D and 4.4E). This difference is likely due to how the two algorithms were developed; while phosphoRS was optimized with both high and low resolution data (Taus *et al.*, 2011), PTM-score was originally developed for phosphosite localization using low mass accuracy ion-trap generated CID data (Cox *et al.*, 2011). Unlike phosphoRS, PTM-score treats all observed MS2 peaks as integer masses (Taus *et al.*, 2011; Chalkley & Clauser, 2012), meaning that there is limited benefit using PTM-score when high resolution data has been acquired. Furthermore, while PTM-score searches the "n" most intense peaks within a bin of 100 *m/z* to identify site-determining product ions, *ptmRS* considers the total number of extracted peaks across the full mass range of the MS2 spectrum, overcoming potential issues of uneven peak distribution in individual *m/z* bins (Taus *et al.*, 2011; Chalkley & Clauser, 2012), and is thus better suited for data generated with high resolution mass analysers.

Both localization tools underestimated the true FLR for EThcD IT data (Fig. 4.4B and C), demonstrating the additional benefit of generating site determining *c/z* as well as *b/y* ions within a single spectrum. A 1% FLR could not be computed for the EThcD OT dataset, as insufficient incorrectly localized phosphopeptides were identified from the library. Instead, the scores defined for this fragmentation mode (PTM-Score = 0.9; *ptmRS* = 99.4, Figure 4.4B, C) represent an FLR of 0.8%. The other PTM-score and *ptmRS* values computed for phosphosite localization at a 1% FLR are broadly in agreement with those previously defined for a larger synthetic phosphopeptide library using a different orbitrap-based MS platform, demonstrating that the MS acquisition methods and the associated bioinformatics platforms are largely transferable between similar platforms (Marx *et al.*, 2013).

In addition to the 1% FDR filtering, 'default settings' in Andromeda apply a score cut off of 40 for post-translationally modified peptides. To investigate whether this artificially reduced the numbers of phosphopeptides identified from our library, all eight datasets were searched again with Andromeda, having removed the requirement for scores to exceed 40 (Supplementary Table 2). An analogous threshold for comparison with MASCOT could not be set since there is not a perfect linear relationship between the two scoring algorithms (Cox *et al.*, 2011). Upon removal of this score filter in Andromeda, the numbers of confidently identified phosphorylation sites was broadly similar, with the exception of the high resolution HCD OT and HCD OT nl EThcD datasets, where an additional seven and six phosphosites were identified respectively. The resultant minimal change in confidently assigned phosphosites (max. 4% with HCD OT; 2% decrease with EThcD IT) meant that amendment of the default settings in Andromeda did not warrant further investigation. Default settings for both search engines were thus used in subsequent investigations, these also being the parameters that most end-users will typically apply.

4.2.5. **Phosphopeptide identification from a phosphopeptide enriched complex human cell lysate**

Having evaluated the eight MS acquisition methods using the phosphopeptide library, six methods for this tribrid MS platform worthy of further investigation based on the numbers of correctly site localized phosphopeptides were defined. Performance of these six MS acquisition strategies for phosphopeptide identification and phosphosite localization was then evaluated using a larger dataset derived from a more complex, biologically relevant sample. Phosphopeptides were enriched from a U2OS cell lysate using TiO_2, and aliquots (6 µl, equivalent to 100 µg from 4 mg digested cell lysate) of the same phosphopeptide enriched sample were analysed in duplicate by LC-MS/MS using HCD OT, HCD IT, EThcD OT, EThcD IT, HCD OT nl EThcD IT or HCD IT nl EThcD IT (Table S1).

The number and overlap of unique phosphopeptide identifications using either Andromeda/PTM-score or MASCOT/*ptmRS* is presented for each of the MS acquisition methods (Table 4.4, Figure 4.5, Appendix 4 Figures S2, S3). Of the six methods assessed, HCD IT exhibited the least overlap between technical replicates, with up to 44% of phosphopeptides being unique to a single LC-MS/MS run. Other methods exhibited between ~80% (HCD OT nl EThcD IT) and 75% (HCD OT) overlap (Figure S1).

The highest total number of unique phosphopeptides from the enriched U2OS cell lysate (6877 phosphopeptides above a 1% FDR) was identified using HCD IT and Andromeda (Table 4.4, Figure 4.5, Appendix Figures S4, S5, S6). This regime maximizes on the capability of the Orbitrap Fusion to parallelize high resolution MS1 acquisition in the orbitrap whilst simultaneously acquiring MS2 data in the ion trap. Interestingly, there was little difference in the numbers of unique phosphopeptides identified using MASCOT when MS2 was performed in the OT *versus* the IT; 4957 phosphopeptides were confidently identified for HCD OT compared with 4920 phosphopeptides using HCD IT (Table 4.4). This is almost certainly due to the enhanced confidence in phosphopeptide identification that results when MS data are acquired with higher mass accuracy, as is the case with HCD OT. However, it is particularly interesting to note how Andromeda and MASCOT differentially handle high resolution and low resolution MS2 data (discussed in more detail below).

Table 4.4. MS acquisition and data analysis methods evaluated using phosphopeptide enriched human cell lysate. For each of the six Orbitrap Fusion MS acquisition methods (Table 1, Table S1) the number of peptide spectrum matches (PSMs) at 1% FDR are presented together with the total number of unique phosphopeptides and phosphosites using either Andromeda with PTM-score (top) or MASCOT and ptmRS (bottom). The number of phosphosites with an FLR ≤ 1% is also presented. [a]Mean values are ± S.D., n = 2.

Search Engine		HCD OT	HCD IT	EThcD OT	EThcD IT	HCD OT nl EThcD	HCD IT nl EThcD
Andromeda/ PTM-score	**# Unique phospho PSMs**[a]	5414 ± 197	6396 ± 728	1947 ± 1	3321 ± 42	5452 ± 88	5506 ± 659
	# unique phosphopeptides	4214	5632	1730	3315	3702	3494
	# phosphosites	4808	6877	1928	3995	4345	4145
	# phosphosites ≤ 1% FLR	2422 (50%)	2550 (37%)	1468 (76%)	3037 (76%)	2472 (57%)	2045 (49%)
MASCOT/ *ptm*RS	**# Unique phospho PSMs**[a]	5118 ± 45	4705 ± 269	2084 ± 26	2847 ± 197	4966 ± 45	4297 ± 69
	# unique phosphopeptides	4957	4920	2148	2947	4153	3398
	# phosphosites	5733	5501	2413	3409	4880	3933
	# phosphosites ≤ 1% FLR	4337 (76%)	3294 (60%)	2078 (86%)	2841 (83%)	3837 (79%)	2717 (69%)

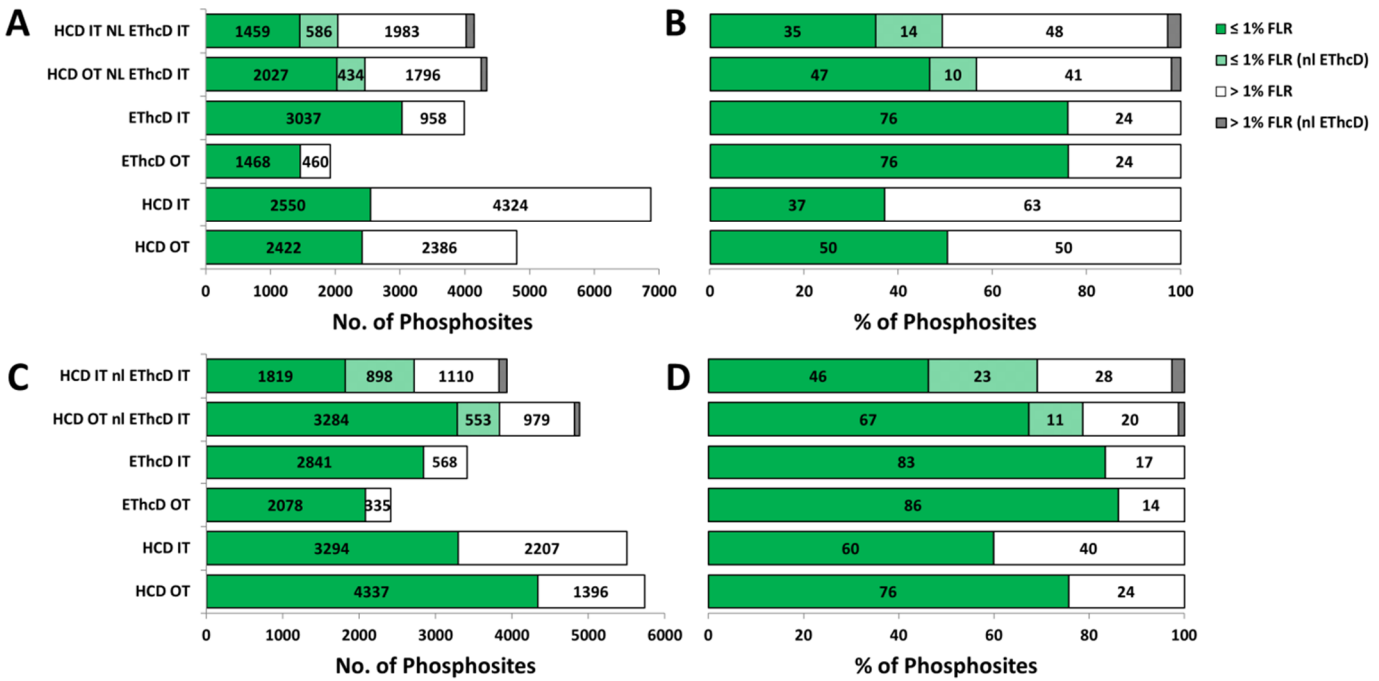


Figure 4.5. Comparison of method-dependent phosphorylation site localization. Confidently localized phosphorylation sites (FLR ≤1%, green) or ambiguous phosphosite assignments (white, grey) from a TiO_2-enriched U2OS cell lysate, using either (A, B) Andromeda/PTM-score, or (C, D) MASCOT/ptmRS for each of the six Orbitrap Fusion MS acquisition methods. Phosphosites assigned by virtue of neutral loss (NL)-triggered EThcD are also presented. Number (A, C) or percentage (B, D) of phosphosites identified is indicated for each condition.

An important reason for undertaking this study was to evaluate confidence in phosphosite localization. Under the conditions examined, phosphosite localization was optimal when utilizing HCD OT and MASCOT/*ptmRS* searching. Of the 5733 phosphosites identified, 76% (4337) were confidently site localized under these conditions (Table 4.4, Figure 4.5, Figures S4, S5, S6). For the same dataset, 4808 phosphosites were defined using Andromeda/PTM-score, of which 50% failed to meet the 1% FLR cut-off for confident site localization using the previously defined PTM-score of 0.994. Although the proportion of confidently site localized phosphopeptides is optimal overall with the EThcD regimes (both OT and IT), as was observed with the phosphopeptide library dataset, the numbers of phosphosites was compromised compared with either the equivalent HCD method, or the neutral-loss driven strategies. Even considering that the site localization scores applied to the EThcD OT data was slightly more conservative (equating to 0.7% FLR, rather than 1% FLR), the distribution of phosphosite localization scores demonstrates that total numbers of phosphosites is still significantly lower with this MS2 method, irrespective of search engine (Fig. S4). Not surprisingly, site localization confidence generally decreased as the number of phosphorylation sites per peptide increased, irrespective of the search algorithm employed (Figures S5, S6). The exception was EThcD OT: ~76% of phosphosites were confidently localized with PTM-score independent of the number of phosphate groups; doubly phosphorylated peptides yielded a higher number of confidently localized phosphosites on average (93%) with *ptmRS* site than singly (86%) or triply (83%) phosphorylated peptides. The performance of Andromeda/PTM-score was uniformly weaker across all datasets compared with MASCOT/*ptmRS*. The exception was the EThcD IT data for singly phosphorylated peptides, where the percentage of confidently localized sites was more comparable for the two search algorithms (78% for Andromeda/PTM-score, 83% for MASCOT/*ptmRS*).

Although the trend in confident phosphosite identification is similar to that observed for the phosphopeptide library, the proportion of incorrect or ambiguous assignments is much higher in the lysate-derived peptides, possibly due the greater diversity of peptide size, and the true/false nature of the manner that the phosphopeptide library was used to define correct/incorrect site localization. In contrast, Andromeda/PTM-score performed much better than MASCOT/*ptmRS* with EThcD IT (but not EThcD OT) data, identifying 12.5% more phosphopeptides, and ~7% more phosphosites with confidence (Table 4.4, Figure 4.5).

For both the HCD OT and HCD IT regimes where nl EThcD IT is triggered, the percentage of confidently assigned phosphosites increases with Andromeda/PTM-score compared to HCD alone, particularly for HCD IT. This reflects the high performance of Andromeda/PTM-score with EThcD IT data. However, the total numbers of phosphosites identified with HCD IT are much lower when neutral loss EThcD is triggered due to the increased time required for ETD. Interestingly, although 42% of HCD IT spectra contained precursor neutral loss product ions (either 98 or 80 amu, at ≥10% base peak signal), a significant number of these were not within the top 10 ions that triggered EThcD, and only 16% of HCD IT spectra precipitated the acquisition of EThcD.

The high proportion of confidently localized phosphosites with EThcD IT (76% and 83% from Andromeda/PTM-score and MASCOT/*ptmRS* respectively), combined with the fact that the two data analysis platforms yielded a high proportion of algorithm unique identifications (Figure 4.6) suggests that this mixed mode fragmentation regime would likely benefit from data interrogation using multiple informatics pipelines: 31% of Andromeda/PTM-score identifications were unique, while 23% were unique to MASCOT/*ptmRS*. Perhaps not unexpectedly, the utility of EThcD OT for high-throughput phosphosite identification was severely compromised due to the additional time required for both ETD and OT-based product ion analysis, resulting in much slower overall acquisition speeds for this high resolution mixed mode fragmentation method. Consequently, there was a ~40-50% decrease in the numbers of confidently localized phosphosites using EThcD OT compared to HCD OT.

The difference in site localization confidence for HCD IT versus HCD OT data for the two algorithms becomes much more apparent for the complex cell lysate derived phosphopeptide sample compared to the synthetic phosphopeptide library, with site localization confidence decreasing from 76% to 60% for MASCOT/*ptmRS* and 50% to 37% for Andromeda/PTM-score (Figure 4.5B, D), again emphasizing the benefits of high resolution MS2 over the reduction in duty cycle afforded by analysis in the ion trap.

Evaluation of the distribution of site localization scores for all phosphopeptides facilitates a better understanding of how the two site localization algorithms handle the different fragmentation modes for this complex phosphopeptide sample (Figure S4). Scoring of EThcD IT data, particularly with *ptmRS*, yields a much shallower distribution of scores than

those for HCD IT. Consequently, large changes in score result in relatively small changes in the number of confidently localized phosphosites. The distribution of scores for HCD OT data is notably distinct between the two algorithms. The elevated mass accuracy of the orbitrap allows *ptmRS* to maximize its ability to pinpoint the correct site of modification, with ~4000 phosphosites having a *ptmRS* score of 100. In contrast, PTM-score consistently scores low resolution ion trap data higher, where the increased ion current and enhanced duty cycle likely yields benefits that are not compensated by the inability of this scoring system to handle high resolution data.

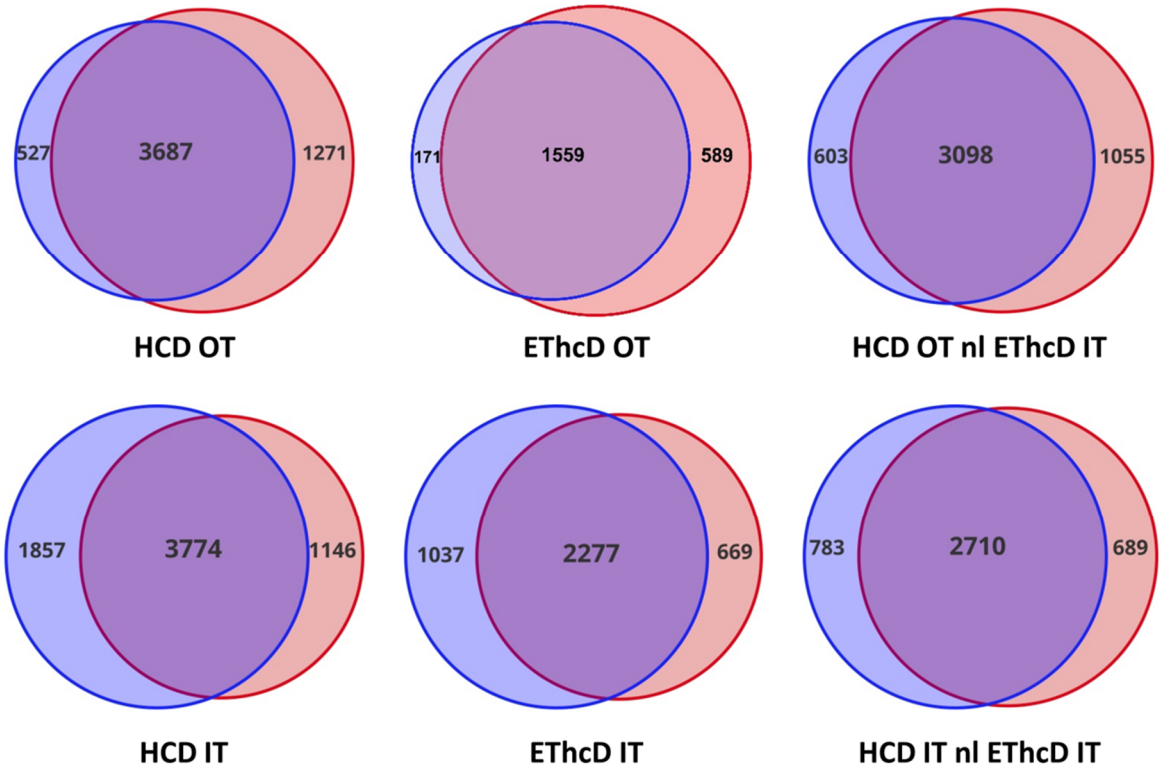


Figure 4.6. Overlap of phosphopeptide identification between search engines. Venn diagrams showing the number and overlap of phosphopeptides identified with either Andromeda/PTM-score (blue, left) or MASCOT/ptmRS (red, right) for each of the six MS acquisition methods applied to TiO_2-enriched U2OS cell lysate**.**

4.2.6. **Confident phosphosite localization is dependent on the number of potential sites of phosphorylation**

To avoid potential confusion when examining the effect of multiple potential sites of phosphorylation (Ser, Thr or Tyr) within a single peptide on site localization confidence, singly phosphorylated peptides only were considered for investigation (Figure 4.7; Figures S7, S8). Unsurprisingly, as the number of Ser/Thr/Tyr residues increases, *i.e.* the number of potential sites of modification increases, the numbers of confidently site localized phosphopeptides decreases with both *ptmRS* and PTM-score. For HCD OT generated tandem mass spectra, this decrease in confident phosphosite localization is much more apparent with PTM-score than with *ptmRS*. For those phosphopeptides containing two Ser/Thr/Tyr residues, the phosphosite is confidently localized in 92% of cases using *ptmRS*, while only 72% are correctly localized with PTM-score. This decreases to 39% for PTM-score when a peptide contains four Ser/Thr/Tyr residues, but only 73% for the same cohort when searched using *ptmRS*. The trend is consistent for HCD OT incorporating neutral loss triggered EThcD, with 80% of the peptides containing 4 Ser/Thr/Tyr residues from the *ptmRS* search having confident site localization, but only 49% being confidently localized by PTM-score (Figure 4.7; Figures S7, S8). For both scoring algorithms, the numbers of confidently assigned sites with HCD OT nl EThcD IT was intermediary between the numbers observed with either HCD OT and EThcD IT, showing potential benefit of the dual fragmentation approach when considering peptides with multiple possible sites of phosphorylation. Under all tandem MS conditions examined, MASCOT/ptmRS performed as well as, or better than Andromeda/PTM-score for confident site localization, irrespective of the number putative sites of phosphorylation (Figure 4.7, Figures S7, S8).

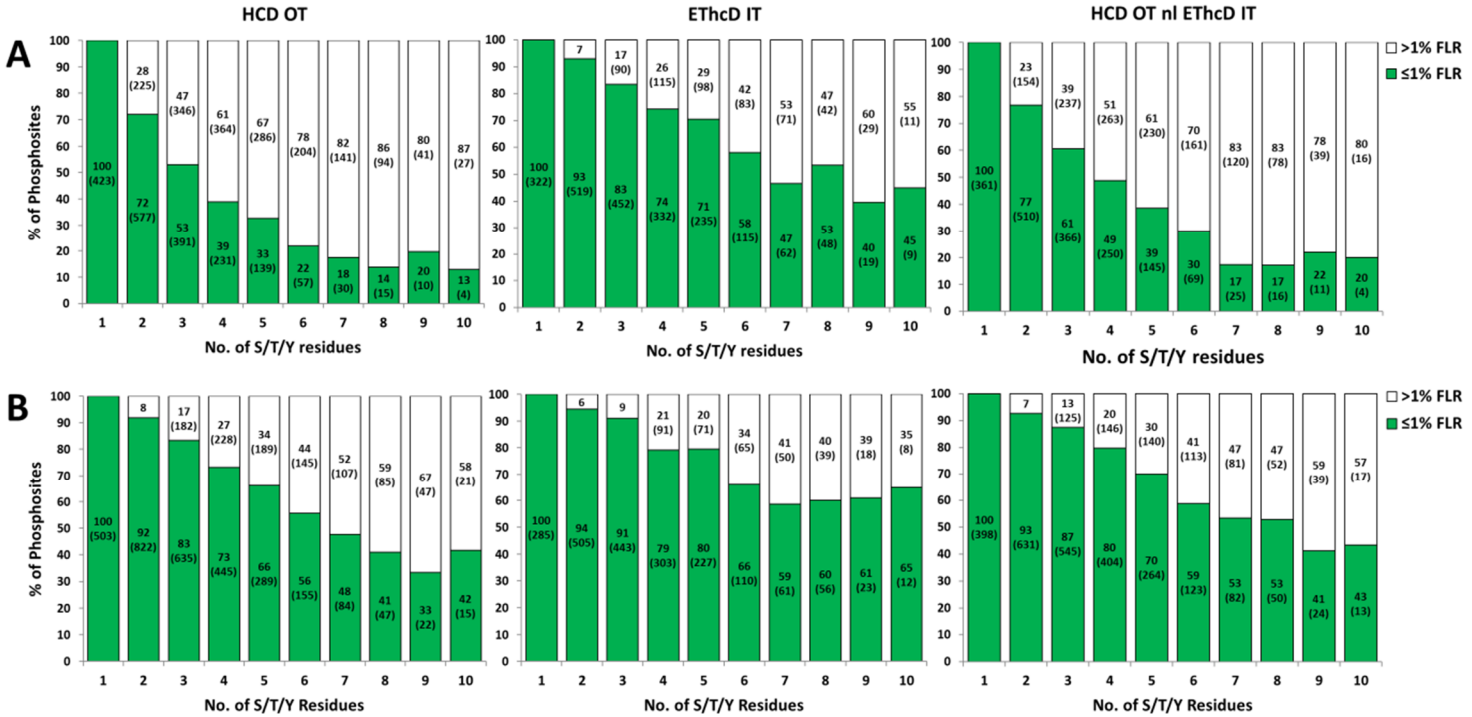


Figure 4.7. Number of confidently localized phosphosites as a function of the number of common putative phosphorylatable residues. Percent correctly site localized phosphopeptides (FLR ≤1%, green) or site ambiguous phosphopeptides (FLR >1%, white) is presented as a function of the number of Ser (S), Thr (T) or Tyr (residues) within the peptide. Data generated by either HCD OT (left), EThcD IT (middle) or HCD nl EThcD OT (right) was search with either (A) Andromeda/PTM-score or (B) MASCOT/ptmRS as previously described. Percentage is indicated for each condition; the number of unique phosphosites is in parentheses. Data for all six MS acquisition methods is presented in Figures S7 and S8.

4.2.7. **Effect of charge state on phosphosite assignment**

It is known that the efficiency of ETD is dependent on charge density and is thus optimal for tryptic peptides in which the charge state is ≥3 (Good *et al.*, 2007). Given that EThcD is a dual fragmentation mechanism, generating both *b/y* (HCD) and *c/z* ions (ETD), the total number of ions generated using this fragmentation regime will thus be dependent on charge state, impacting the number of site-determining product ions. The effect of charge state on phosphosite localization confidence was therefore evaluated (Figure 4.8, Figures S9, S10). Unsurprisingly, the ability to pinpoint the site of modification was notably improved with EThcD IT compared with HCD IT alone for precursor ions where *z*=3, with either 84% (MASCOT/ptmRS) or 75% (Andromeda/PTM-score) of phosphosites being defined by EThcD, compared with 53% or 29% respectively for HCD IT. The same is true for EThcD OT compared with HCD OT, with 77% or 42% respectively of 3+ peptide ions being correctly site localized with PTM-score, *c.f.* 87% (EThcD OT) and 66% (HCD OT) with *ptmRS* (Figs. S9, S10). EThcD IT also outperformed both HCD OT and HCD IT for confident site localization for ions of charge states 2+ and 4+, albeit with significantly fewer phosphosites being identified in total with EThCD IT than with either HCD method for 2+ ions (Figure 4.7, Figs. S9, S10).

Both of the MS acquisition strategies invoking EThcD as a consequence of precursor neutral loss (HCD IT nl EThcD; HCT OT nl EThcD) were compromised in terms of the efficiency and total number of phosphosites identified for 3+ and 4+ ions, with no apparent benefit.

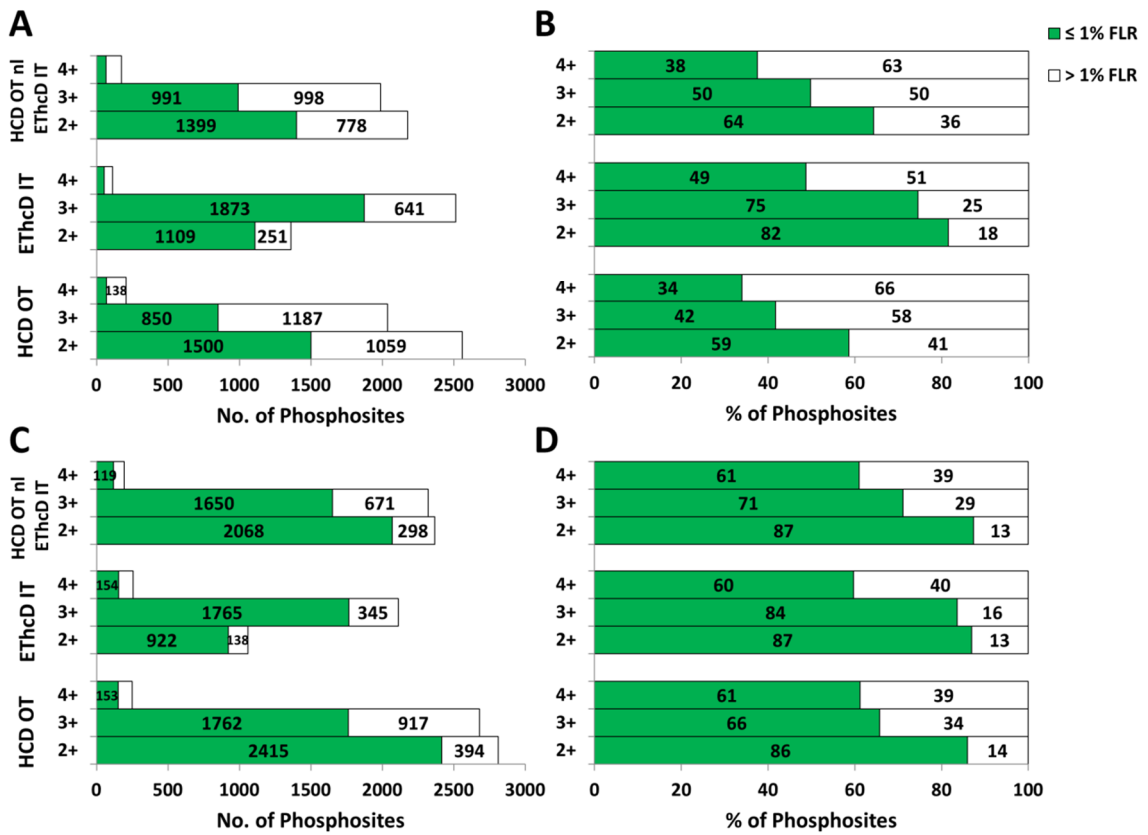


Figure 4. 8. Phosphosite localization as a function of peptide ion charge state. Confidently localized phosphorylation sites (FLR ≤1%, green) or ambiguous phosphosite assignments (FLR >1%, white) presented as a function of precursor ion charge state for data searched with Andromeda/PTM-score (A, B) or MASCOT/ptmRS (C, D). Number (A, C) or percentage (B, D) of phosphosites identified is indicated for each condition. Data for all six MS acquisition methods is presented in Figures S9 and S10.

4.2.8. **Assessment of a charge-dependent HCD/EThcD method**

Following the publication of the 8 acquisition methods and showing the limited benefit of a neutral loss triggered EThcD method for phosphoproteomics studies due to the relatively small number of ions triggered for ETD fragmentation, an additional method was considered with the aim of combining the benefits of both HCD and EThcD. In this charge dependent method, ions with a charge (*z= 2+*) were fragmented with HCD whereas ions with a charge (*z= 3+ to 5+*) were fragmented with EThcD. ETD of 2+ peptide ions is inefficient due to electron transfer being less exothermic for 2+ peptides than 3+ or 4+ ions, and that the singly charged cation formed from a 2+ ion following electron transfer has greater kinetic stability than the doubly/triply charged cation formed from 3+ and 4+ peptide ions. The Coloumb repulsion is therefore absent in the singly charged cation following electron transfer and so non-dissociative electron capture occurs leading to highly inefficient fragmentation (Pitteri *et al.*, 2005; Swaney *et al.*, 2007). However, as the majority of peptides produced by trypsin digestion are doubly charged following electrospray ionization, it is crucial to obtain high quality data for these peptides and so a charge-dependent HCD/EThcD method represents an ideal approach for obtaining high numbers of identifications and maximising confident phosphosite localisation. Similar to the methods described in section 4.3, the HCD fragment ions were detected in the orbitrap and EThcD fragment ions detected in the ion trap. The method was compared against the HCD OT NL EThcD IT method to evaluate the possible benefits of this approach.

4.2.9. **Assessment of phosphopeptide identification & phosphosite confidence**

A phosphopeptide enriched U2OS lysate was analysed in duplicate and revealed 15 % fewer unique phosphopeptides identifications with the charge dependent method compared with the neutral loss triggered method and this is likely due to a higher number of ETD events resulting in a longer reaction times and fewer spectra acquired (Table 4.5).

Table 4. 5. MS acquisition methods evaluated using phosphopeptide enriched human cell lysate. For both methods, the number of peptide spectrum matches (PSMs) at 1% FDR are presented together with the total number of unique phosphopeptides and phosphosites identified using MASCOT with ptmRS. The number of phosphosites with an FLR ≤ 1% is also presented.

Search Engine		HCD OT (2+) EThcD IT (3+ to 5+)	HCD OT nl EThcD IT
MASCOT/ *ptm*RS	# Unique phospho PSMs[a]	4894	3567
	# unique phosphopeptides	2148	2571
	# phosphosites	2579	3046
	# phosphosites ≤ 1% FLR	2116 (82%)	2279 (75%)

When considering phosphosite localisation, 82 % of phosphosites were confidently localised (≤ 1% FLR) with the charge state-dependent regime compared with 75 % for the NL triggered method. This is comparable to the EThcD OT & IT methods described in section 4.3. However, when considering the total number of high confidence phosphosites identified, fewer sites were identified using the charge state-dependent method (2116 *vs* 2279) suggesting additional optimization (*e.g.* ETD reaction times) to increase spectral acquisition would be required to maximize the benefits of this method (Table 4.5; Figure 4.9). With further development, this method could be usefully implemented in phosphopeptide studies requiring EThcD spectra where it is most beneficial (*z* = *3*+) whilst obtaining high quality HCD spectra for 2+ ions.

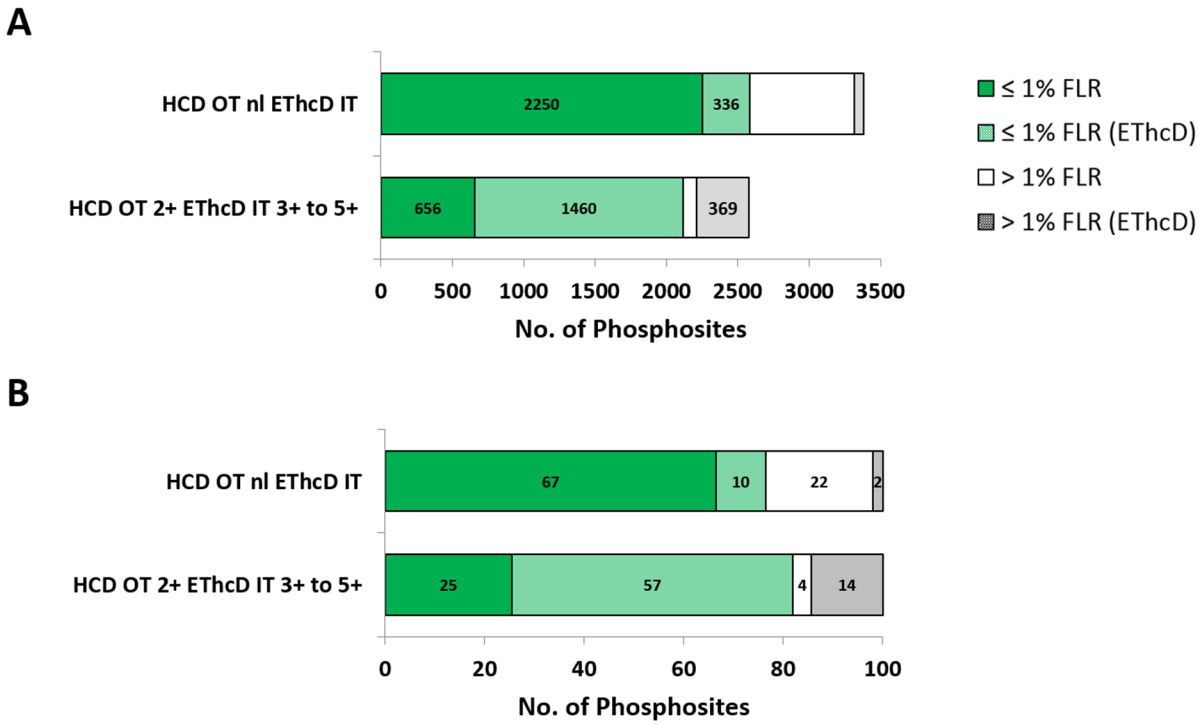


Figure 4.9. Comparison of method-dependent phosphorylation site localization. Confidently localized phosphorylation sites (FLR ≤1%, green) or ambiguous phosphosite assignments (white, grey) from a TiO_2-enriched U2OS cell lysate, using Mascot with ptmRS to process data from both HCD OT nl EThcD IT (A) and HCD OT (2+) EThcD IT (3+ to 5+) acquisition methods. Phosphosites assigned by EThcD are also presented (light green). The number (A) and percentage (B) of phosphosites identified is indicated for both methods tested

4.2.10. **Effect of number of phosphorylatable residues (S/T/Y) on phosphosite assignment**

Next, the ability to localize phosphosites as the number of possible phosphorylatable sites (Ser/Thr/Tyr) increased was assessed. Compared with the NL triggered method, the charge-dependent method identified more sites with high confidence. With 5 possible sites of phosphorylation, the charge-dependent method confidently localized 81 % of phosphosites, compared with 69 % for the NL triggered method and for peptides containing 9 possible phosphosites, the charge dependent method confidently localized 58 %, compared with 40 % for the NL triggered method (Figure 4.10). This is most likely due to the method utilizing HCD and EThcD fragmentation where each is most beneficial (for 2+ ions and >3+ respectively) which increases the quality of the resulting MS2 spectra, allowing for greater confidence in the site of modification. This further highlights the potential benefits of this method for phosphoproteomics studies.

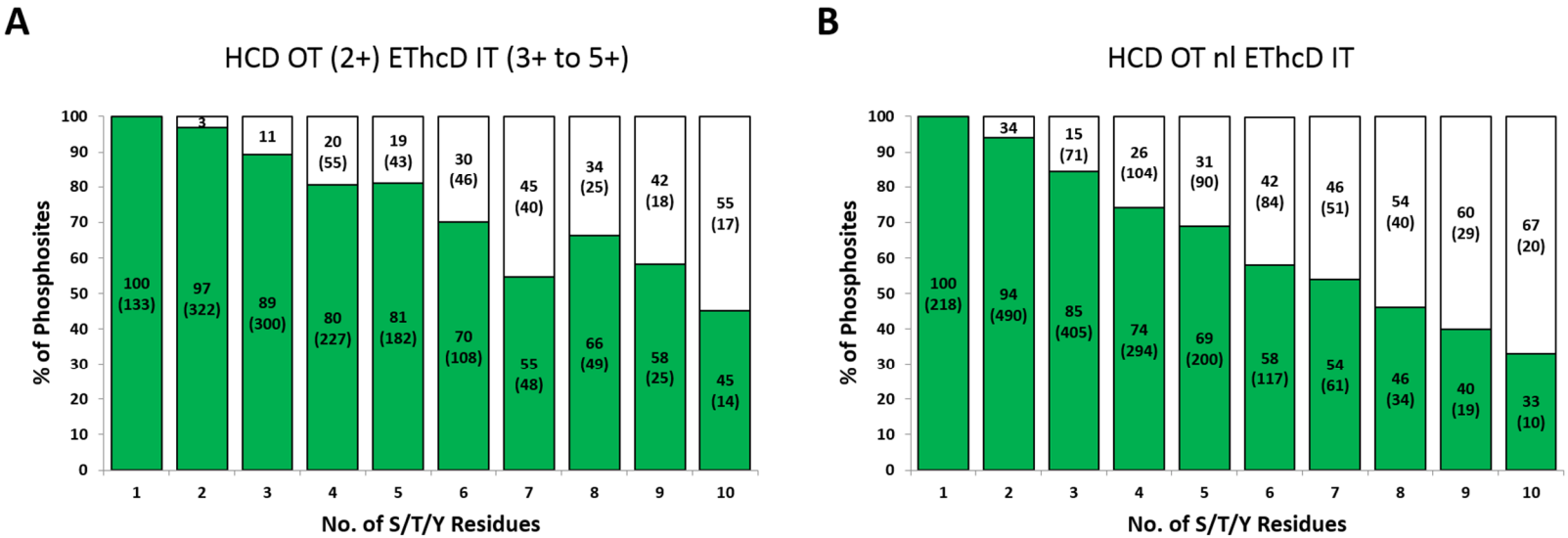


Figure 4.10. Number of confidently localized phosphosites as a function of the number of common putative phosphorylatable residues. Percent correctly site localized phosphopeptides (FLR ≤1%, green) or site ambiguous phosphopeptides (FLR >1%, white) is presented as a function of the number of Ser (S), Thr (T) or Tyr (residues) within the peptide. Data generated by either HCD OT (2+) EThcD IT (3+ to 5+) (A) or HCD OT nl EThcD IT (B) was searched using Mascot with ptmRS and the percentage is indicated for each condition; the number of unique phosphosites is in parentheses.

4.2.11. **Effect of charge state on phosphosite assignment**

Similarly, the charge-dependent method also outperformed the NL triggered method when considering phosphosite localization of peptides as a function of charge state. There was little difference between the percentage of confidently assigned phosphosites for 2+ ions (87% for the charge dependent method vs 85 % for the NL triggered method) which is expected, as the majority of 2+ ions in the NL method, and all in the charge dependent method were identified with HCD, rather than EThcD (Figure 4.11).

For 3+ peptide ions, the charge dependent method outperformed NL triggered method, with 84 % confidently localized compared with 73 % (Figure 4.11). This is due to all 3+ peptides being fragmented with EThcD using this method, allowing for high quality MS2 spectra to be acquired. However, higher total numbers of phosphosites for >3+ peptide ions were identified with the NL method and this again highlights the requirement of further development of the charge dependent method to utilize cycle time most efficiently to acquire higher numbers of high quality MS2 spectra.

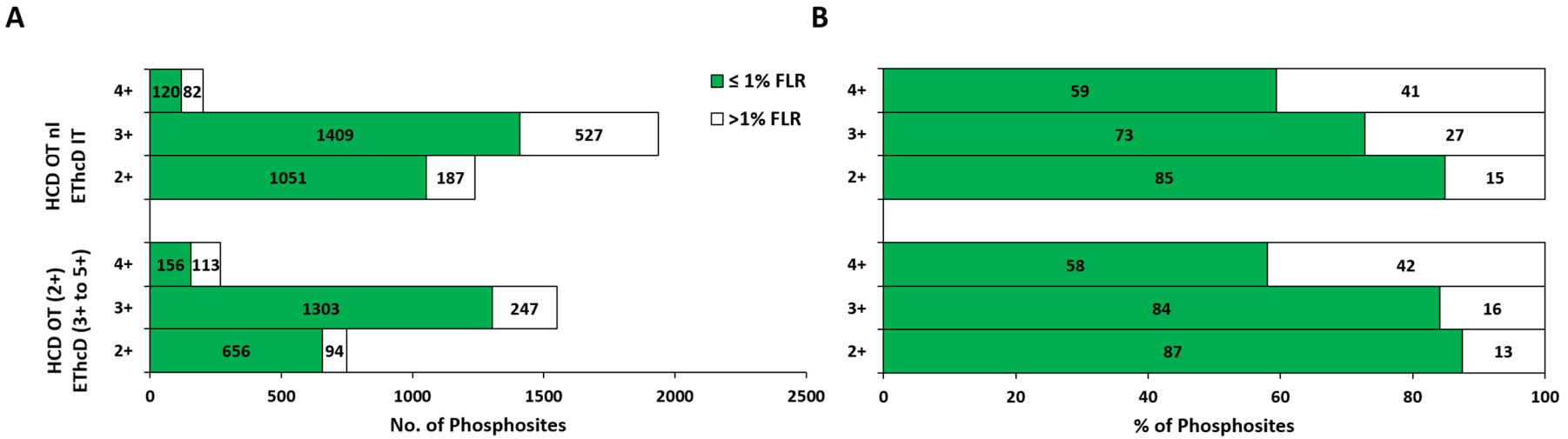


Figure 4.11. Phosphosite localization as a function of peptide ion charge state. Confidently localized phosphorylation sites (FLR ≤1%, green) or ambiguous phosphosite assignments (FLR >1%, white/grey) presented as a function of precursor ion charge state for data searched MASCOT with ptmRS. The number (A) and percentage (B) of phosphosites identified is indicated for both methods tested**.**

Taken together, the charge dependent method for phosphoproteomics studies represents an ideal method for confidently identifying sites of phosphorylation with EThcD fragmentation where most beneficial (z= 3+ to 5), and HCD for 2+ ions, to obtain a balance between number of identifications and confident phosphosite localisation. Similar to the methods described in section 4.3, the total numbers of identifications decrease due to a greater number of ETD events compared with the NL triggered methods and therefore, this method is likely to be most beneficial for studies requiring ETD fragmentation whilst aiming to increase identifications with HCD fragmentation for 2+ peptide ions.

4.2.12. **Conclusions**

The work described in this chapter aimed to develop and optimize MS acquisition methods for large-scale phosphoproteomics studies and assess downstream processing tools to maximize phosphopeptide identifications with confident phosphosite localization.

An assessment of both collision energy and ion injection time for MS2 led to the optimization of these parameters for the analysis of phosphopeptides. Eight MS acquisition strategies on the Orbitrap Fusion mass spectrometer were subsequently tested, for their ability to confidently identify and, crucially, to pinpoint sites of modification on phosphopeptides. In addition, the relative efficiency of two of the most widely used phosphoproteomics data analysis platforms for optimal phosphosite identification: MASCOT integrated into Proteome Discover using *ptmRS*, and Andromeda with PTM-score was examined.

Using a synthetic phosphopeptide library, MS method-specific scores for Andromeda/PTM-score and MASCOT/*ptmRS* that yielded a 1% FLR were defined. When applied to a complex biologically-derived phosphopeptide mixture, even small changes in the applied scores may yield significant changes in the numbers of phosphosites identified for HCD-mediated fragmentation, and the marked difference in site confidence for the different MS methods at any given value cannot be ignored.

The findings are largely in agreement with previous observations made using other orbitrap-based MS platforms, which demonstrate that phosphosite localization confidence is optimal with EThCD where a dual ion series is generated (Frese *et al.*, 2013). However, the total number of unique phosphopeptides identified, as well as the number of

confidently localized phosphosites, is optimal when high resolution analysis of HCD fragment ions for MS2 is employed. MS acquisition strategies invoking neutral loss-mediated ETD-based fragmentation is hampered by both the additional time taken to perform this type of fragmentation in a second round of MS2, as well as the surprisingly few phosphopeptide ions that generate neutral loss product ions and thereby invoke this second round of MS2 analysis.

Differences in the ways that the two bioinformatics platforms handle distinct types of tandem MS data and the number of unique phosphopeptides identified, means that there is likely to be benefit in searching data acquired using a single acquisition strategy using both data analysis pipelines. This is particularly apparent with EThcD, where 31% and 23% of phosphopeptides respectively are unique to either Andromeda/PTM-score or MASCOT/ptmRS. The relatively few unique phosphopeptide identifications with Andromeda for HCD OT data, and the overall reduction in confident site localization using Andromeda/PTM-score for regimes exploiting fragmentation strategies other than EThcD, means that multi-algorithm searching may not be of significant benefit with other types of data.

Whilst EThcD methods provides the highest confidence in sites of phosphorylation, the total numbers of identifications were much lower due to the long reaction times required for efficient fragmentation and therefore, it was concluded that EThcD methods were unsuitable for a large-scale phosphoproteomics study where high number of identifications are required to gain a global assessment of the phosphoproteome under study. However, future advances in MS platforms and ETD fragmentation will undoubtedly permit the use of ETD for large-scale phosphoproteomics studies.

The data also highlights that there are likely to be additional benefits in terms of increased numbers of confidently localized phosphosites, by implementing EThcD for ions with charge state of $\geq$3+ and the employment of a multiple algorithm search strategy. Moreover, the 'high-definition ETD' (ETD HD) permissible with the Orbitrap Fusion Lumos, which is reported to facilitate ETD on larger precursor ion populations, will likely result in even greater benefits when applied to such a charge-state mediated data acquisition strategy for phosphoproteomics.

Currently, HCD OT methods on the Orbitrap Fusion represent the best approach for large-scale phosphoproteomic studies, with high numbers of phosphopeptides identified with high confidence in the site of phosphorylation. This method will therefore be used for the large-scale phosphoproteomics analysis of the PLK4 regulated phosphoproteome. In addition to this analysis, the unenriched peptides will also be analysed to study protein level changes upon PLK4 inhibition with centrinone. Although the HCD IT method produced the lowest phosphosite confidence, the ability to parallelize MS1 in the orbitrap and, thus, obtaining high resolution data required for quantification with MS2 scans in the ion trap, does, however, provide a method capable of generating high numbers of peptides and this method will therefore be used to study the unenriched proteome. This analysis will be the focus of the next chapter.

5. Investigating Polo-like kinase 4-regulated signalling by SILAC-based quantitative phosphoproteomics

5.1. Introduction

Following the development of new tools for the functional analysis of PLK4 in human cells (described in Chapter 3), and the optimisation of mass spectrometry acquisition methods for phosphopeptide analysis, a large scale (phospho)proteomics study was undertaken to identify PLK4 substrates and PLK4-regulated signalling pathways. Whilst a greater understanding of the function of PLK4 during centriole biogenesis has been achieved in recent years, no global assessment of the roles of PLK4 in the cell, including its full substrate repertoire, has been undertaken thus far.

As discussed in Chapter 1, PLK4 is the master regulator of centriole biogenesis, with the activities of PLK4, STIL & SAS6 being carefully co-ordinated to ensure that each centrosome is duplicated once per cell cycle (Dzhindzhev *et al.*, 2014; Kratz *et al.*, 2015; Moyer *et al.*, 2015). Dysregulation of PLK4 leading to overexpression results in supernumerary centrosomes and mis-segregation of chromosomes, which is a hallmark of many cancers. PLK4 has therefore been identified as a viable target for the development of chemotherapeutics to inhibit protein activity and has led to one compound, CFI-400945, reaching phase I clinical trials (Mason *et al.*, 2014). Therefore a comprehensive insight in to the activities of PLK4 is crucial, to help understand its multiple roles at the centrosome and perhaps wider roles in the cell cycle, so that the potentially complex effects of PLK4 inhibition by compounds such as centrinone can be understood (Wong *et al.*, 2015). A large-scale analysis of the PLK4 regulated (phospho)proteome was therefore deemed a suitable approach to globally assess cellular PLK4 activity.

Global (phospho)proteomics analysis has previously been performed to study the activities of another polo-like kinase family member, PLK1, and revealed regulated phosphosites linked to PLK1 activity via their roles in mitotic progression. In addition, novel substrates were identified, which implicated PLK1 in the DNA damage response and cohesin release from chromosomes (Grosstessner-Hain *et al.*, 2011). Similarly, a global assessment of the Aurora A regulated (phospho)proteome revealed roles in RNA processing, splicing and DNA damage repair in addition to identifying known substrates implicated in mitosis

(Kettenbach *et al.*, 2011b). This further highlights the benefits of a discovery MS approach to study global phosphorylation events, as this is the only method capable of analysing the full range of activities of a particular protein kinase in an untargeted manner.

The results presented in this chapter represent the first global assessment of PLK4 signalling in human cells, with the identification of a number of PLK4-regulated proteins and phosphosites. The regulated sites were mapped on to known pathways, implicating PLK4 in activities beyond that of centriole biogenesis.

5.1.1. **Aims**

The aims of the work in this chapter were to perform a high-throughput SILAC-based quantitative phosphoproteomics investigation to uncover PLK4 regulated signalling pathways. This involved the analysis of both the total proteome and enriched phosphopeptides to assess changes in protein expression and phosphorylation state in response to PLK4 inhibition with centrinone. This work also attempted to validate potential PLK4 substrates through the parallel analysis of a 'drug-resistant' G95R PLK4 allele, which could be inducibly expressed in stable cells treated with and without centrinone. Finally, this work aimed to validate possible PLK4 substrates by western blotting, and generate new information relevant to cellular PLK4 substrate specificity, which has only previously been analysed *in vitro* relative to the PLK1-3 consensus motif (Johnson *et al.*, 2007; Leung *et al.*, 2007).

5.2. Results & Discussion

5.2.1. **Strategy for the phosphoproteomics analysis of PLK4 signalling**

U2OS cells stably transfected with either FLAG-WT or G95R PLK4 (developed in Chapter 3) were used for the work described in this chapter. Cells grown in light (Arg0 Lys0) or heavy (Arg10 Lys8) media were treated with DMSO and centrinone respectively (Figure 5.1). Four bioreplicates were prepared (following separate cell expansions from the same aliquot of cells) and analysed by western blotting, which confirmed the successful overexpression of WT and G95R PLK4 in all four bioreplicates. An accumulation of WT PLK4 is observed following treatment with centrinone, reflecting the inhibition of PLK4 activity, preventing trans-autophosphorylation within the degradation motif to signal for proteasomal degradation. No accumulation of G95R PLK4 was observed, confirming the drug-resistant phenotype (Figure 5.2).

The cell lysates were mixed 1:1, digested with trypsin and fractionated by high pH reversed phase chromatography. Fractions were concatenated in to 12 pools and an aliquot of each (5 µg) was removed for total proteomics analysis using the HCD IT acquisition method on the Orbitrap Fusion described in Chapter 4. The remaining samples were phosphopeptide enriched and analysed using the HCD OT method. Following MS acquisition, the data were processed in MaxQuant. The full workflow is shown in Figure 5.1.

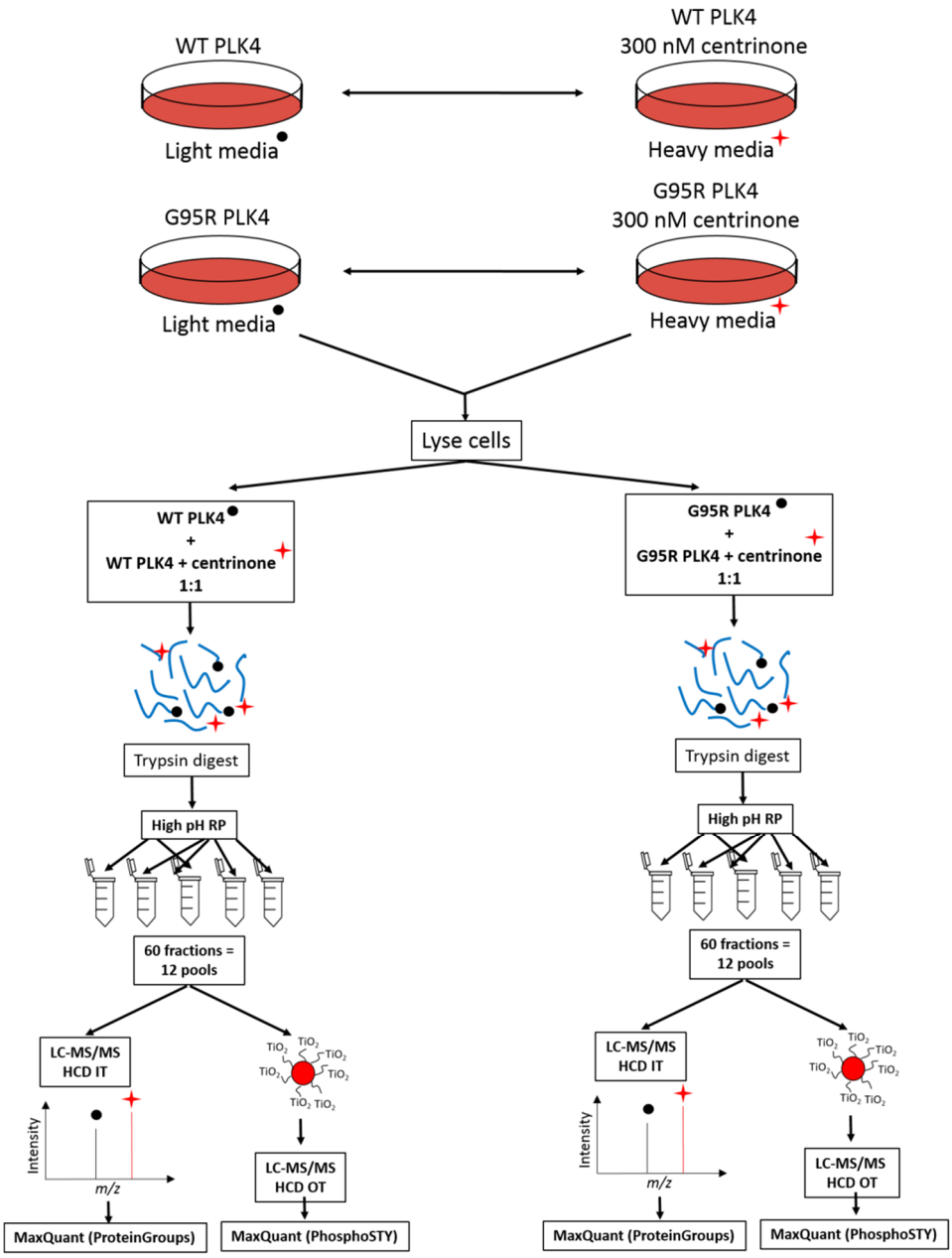


Figure 5.1. Workflow for SILAC-based quantitative (phospho)proteomics. FLAG-WT and G95R PLK4 cells were grown in light (R0K0) and heavy (R10K8) media. Light cells were treated with DMSO and heavy cells were treated with 300 nM centrinone for four hours. Cells were lysed and heavy and light samples were mixed 1:1. Lysates were digested with trypsin and fractionated by high pH reversed phase chromatography. The 60 fractions collected were concatenated to generate 12 pools. Aliquots were removed for analysis of the total proteome by LC-MS/MS and the remaining sample was phosphopeptide enriched. Data were processed in MaxQuant.

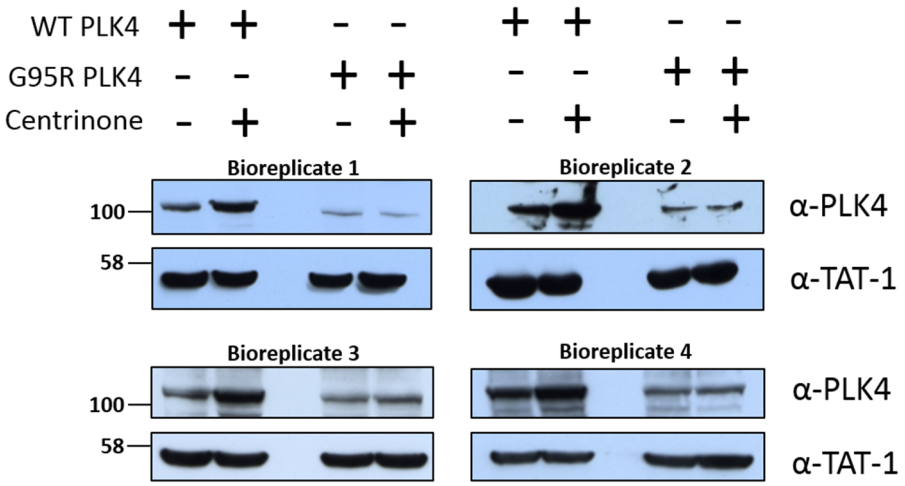


Figure 5.2. Analysis of overexpressed and centrinone-treated FLAG-WT & G95R U2OS cells. Expression of PLK4 was induced upon incubation with 1 µg/mL tetracycline for 18 hours. Cells were treated with 300 nM centrinone or 0.1 % (v/v) DMSO control for 4 hours. Total lysates were analysed by western blot using the indicated antibodies.

Whilst SILAC is a robust method for quantification, labelling efficiency can be reduced by the metabolic conversion of arginine to proline during cell culture. This conversion greatly affects the accuracy of quantification, as it lowers the intensity of the heavy labelled peak and introduces additional satellite peaks containing heavy proline that will not be considered for quantification (Ong *et al.*, 2003; Blagoev & Mann, 2006). It is therefore important to assess arginine to proline conversion before proceeding with statistical analysis of the acquired data. Figure 5.3 confirms that no arginine to proline conversion occurred during metabolic labelling of cells as the density remains centred on the Log2 H/L ratio of 0, and no drift toward the light peptide is observed as the number of proline residues increases. This provided confidence that H/L ratios had been accurately quantified and could therefore be used to determine significantly changing proteins and phosphosites.

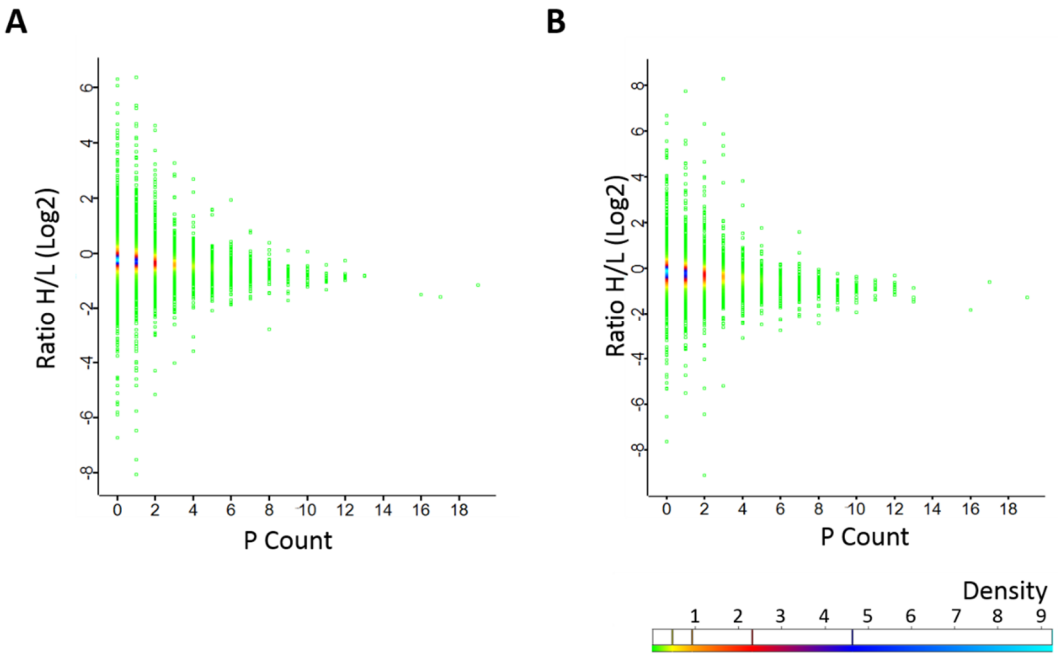


Figure 5.3. Assessment of Arg to Pro conversion. Density plots were generated to assess the metabolic conversion of Arg to Pro in WT PLK4 (A) and G95R PLK4 (B) expressing cells. Non-normalised H/L ratios were plotted against the total proline count from the identified peptides. The data points are colour coded based on density. No global drift toward the light peptide was observed in either cell line, confirming that metabolic conversion had not occurred.

An initial assessment of the overlap between the four bioreplicates was performed to check reproducibility. A high degree of overlap was obtained for protein identifications, with 96 % of proteins identified in ≥3 WT bioreplicates and 94 % of proteins identified in ≥3 G95R bioreplicates. Much greater variability was observed for the phosphopeptide data, which is expected due to the stochastic nature of peptide identifications when data is acquired using a DDA approach. A total of 13,816 phosphopeptides were identified in the WT dataset, with 54 % identified in ≥3 bioreplicates. In the G95R dataset, a total of 13,060 phosphopeptides were identified, with 48 % identified in ≥3 bioreplicates (Figure 5.4).

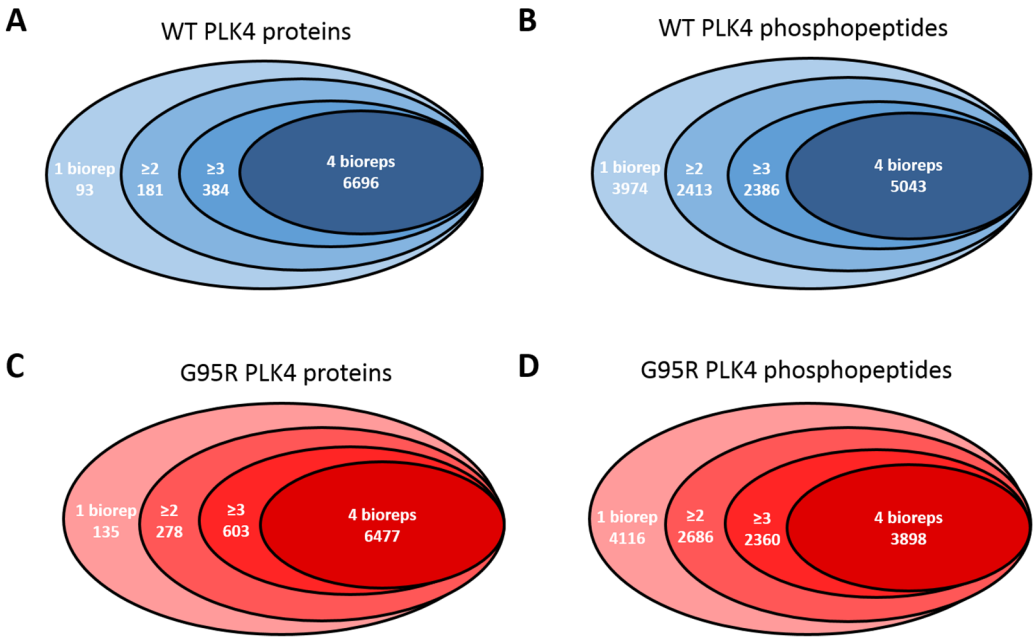


Figure 5.4. Overlap between biological replicates. The overlap between proteins (A, C) and phosphopeptides (B, D) identified between the four bioreplicates for both WT and G95R PLK4 overexpressing U2OS cells.

5.2.2. **Assessment of protein expression**

Statistical analysis of the H/L ratios of proteins quantified in ≥3 bioreplicates was performed using an empirical Bayesian approach; with the Benjamini-Hochberg procedure used for multiple comparisons correction to generate FDR adjusted *p* values (Kammers *et al.*, 2015). Figure 5.5 A & B show the Log2 fold changes of proteins quantified in WT and G95R PLK4 cell lines plotted against adj. *p* value. Proteins with adj. *p* value ≤0.05 are labelled and highlighted in red. In the WT PLK4 dataset, 87 proteins were identified as significantly changing, with 67 downregulated upon centrinone treatment, and 20 upregulated. WT PLK4 was identified with a H/L ratio of 1.8 and an adj. *p* value of 0.052. The H/L ratio reflects the accumulation of PLK4 in the cell upon inhibition with centrinone and corroborates the data in Figure 5.2. In the G95R dataset, only one PLK4 peptide was identified, in one bioreplicate and therefore the protein was not quantified. This could be due to the lower expression of G95R PLK4 compared to WT, as shown by western blot analysis in Figure 5.2. Compared with the WT dataset, a much greater number of proteins were identified as significantly regulated in the G95R PLK4 expressing cells. In total, of the 7492 proteins identified in this dataset, 5949 were quantified, and 1791 proteins were identified as significantly changing, with 1240 downregulated and 551 upregulated

following treatment of the cells with centrinone (adj. *p* value ≤0.05) (Figure 5.5 B; Figure 5.6 B).

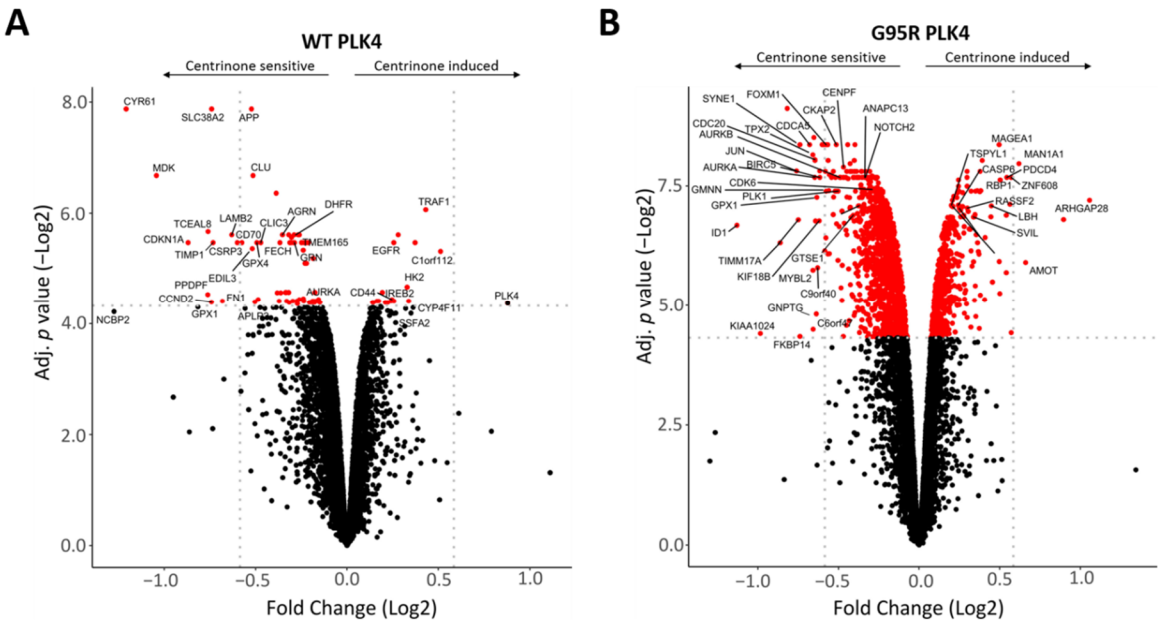


Figure 5.5. Proteomics analysis of centrinone treated WT- and G95R PLK4 expressing cells. H/L ratios were tested for significance using a Bayesian statistical approach. Log2 fold change of identified proteins was plotted against –Log2 Benjamini-Hochberg FDR adj. *p* values for both the WT (A) and G95R (B) PLK4 datasets. Proteins with an adj. *p* value ≤0.05 are highlighted in red.

Functional annotation and enrichment analysis of the significantly regulated proteins in the G95R PLK4 dataset was performed using DAVID. Figure 5.6 A shows an enrichment of dysregulated proteins localised at mitochondrion, focal adhesions and within the cell-cell adherens junctions. In addition, an enrichment of proteins involved in the regulation of metabolic pathways, amino acid biosynthesis and ribosome biogenesis was observed, which would suggest a cellular response to stress upon inhibition with centrinone (Tennant *et al.*, 2010; Cairns *et al.*, 2011; Poliakova *et al.*, 2018). Dysregulated proteins were also implicated in activities at later stages of the cell cycle, such as cytokinesis and metaphase plate congression, whereby chromosomes are aligned at the metaphase plate. Mitotic proteins including Aurora A, TPX2, cyclin B1, and FOXM1 were all identified as significantly downregulated upon treatment with centrinone. Functional network analysis using STRING revealed the connection between these downregulated G2/M phase proteins (Figure 5.6 C), suggesting the possibility that centrinone treatment of this cell line resulted in G2/M phase arrest.

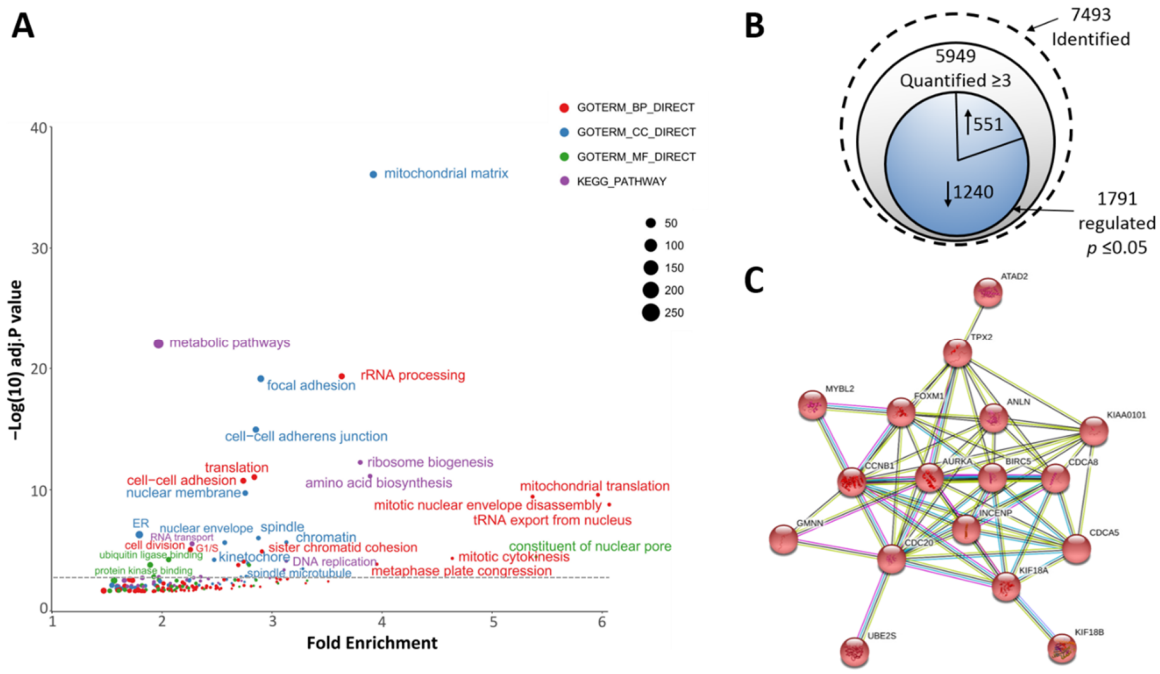


Figure 5.6. Enrichment and network analysis of significantly regulated proteins in the G95R PLK4 dataset. A) GO term enrichment analysis of significantly regulated proteins. Proteins enriched with an FDR adj. p value ≤0.05 (using the Benjamini-Hochberg procedure) are shown above the dotted line and labelled. BP= biological processes; CC = cellular compartment; MF= molecular function. B) The total number of proteins identified in the G95R PLK4 dataset is shown, along with the total number quantified and proteins identified as significantly regulated ($p \leq 0.05$). C) STRING interaction analysis revealed a network of mitotic proteins identified as downregulated.

5.2.3. **Cell cycle analysis**

To test whether centrinone had led to mitotic arrest in the G95R PLK4 cells, U2OS cells were incubated +/- tetracycline to induce WT or G95R PLK4 expression. The cells were then incubated +/- centrinone for four hours, to match the conditions used for the phosphoproteomics study and the cell cycle distribution for each condition was assessed by flow cytometry. Figure 5.7 shows the percentage of cells in G0/G1, S, or G2/M. The overall cell cycle profile for both WT and G95R cells is consistent with the reported profile for asynchronous U2OS cells (Di *et al.*, 2014). The observed shift in profile between WT and G95R cells may be due to the differential effect on growth rates arising from expression of exogenous proteins in this U2OS cell system.

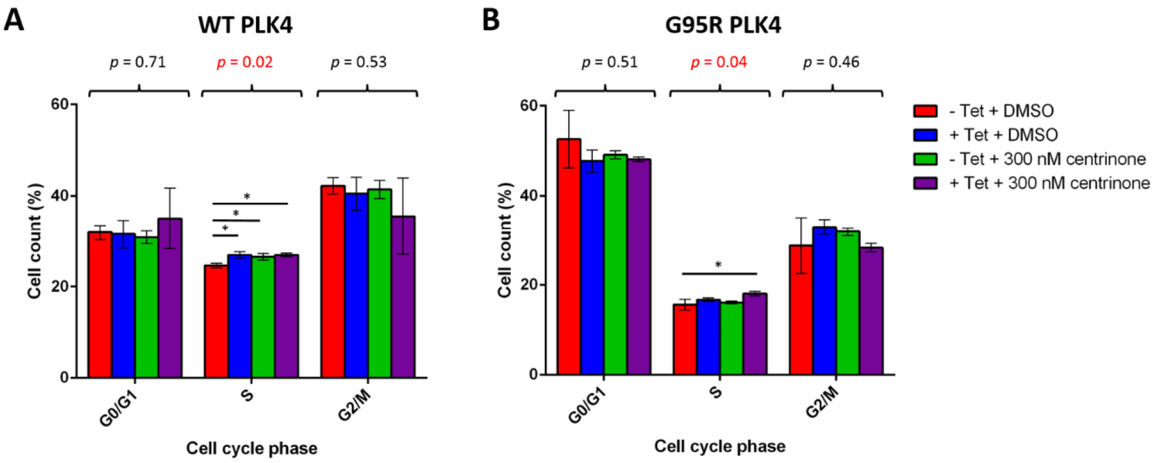


Figure 5.7. Cell cycle analysis by flow cytometry. Expression of PLK4 was induced upon incubation with 1 µg/mL tetracycline for 18 hours. Cells were treated with 300 nM centrinone or 0.1 % DMSO for 4 hours cell cycle distribution was analysed by FACs. The bar chart reflects the percentage of cells in each cell cycle phase for WT (A) and G95R (B). The data is plotted as the mean ±SD for three biological replicates. ANOVA generated *p* values are shown above each phase and the * indicated on the chart represents significant differences between individual samples determined using Tukey post-hoc testing (adj. *p* value ≤0.05).

Statistical analysis using ANOVA revealed no statistically significant differences between cell cycle phases at G1 or G2. However, a significant difference was observed in S phase for both WT and G95R cells. In the WT cells, the induction of PLK4 and the presence of centrinone revealed a statistically significant increase of cells in S phase, compared with all other conditions (Figure 5.7 A). Similarly, G95R PLK4 expressing cells incubated with centrinone revealed a significant increase in S phase compared with the negative control (-Tet/+DMSO) (Figure 5.7 B). A significant increase in the percentage of cells in S phase should be accompanied by a statistically significant difference in the percentage of cells in G1 or G2 phase, however, this was not observed. It is possible that the small difference (~2.5 %) is split between G1 and G2, distributing the effect across both, resulting in neither being determined to be statistically significant. The numbers of cells in S phase also show less variance than the other two cell cycle phases between replicates, which allows smaller differences to be identified as significant. Overall, the results are inconclusive and the experiment would need to be repeated to confirm whether there is a true difference in cell cycle distribution between the different conditions for both cell lines.

The typical cell doubling time of U2OS cells is ~29 hours (Musa, 2013), therefore, the four hour incubation of cells with centrinone was unlikely to be sufficient to induce full mitotic arrest that is observable by cell cycle analysis. The reasonably small ~1.5 fold decrease in

mitotic proteins identified in the STRING network suggests that the centrinone treated G95R PLK4 expressing cells may be at the onset of mitotic arrest. However, cell cycle profiling following longer centrinone incubation times would be required to confirm this.

5.2.4. **Assessment of WT PLK4 regulated proteins**

The comparatively fewer proteins exhibiting differential expression (1 % of total) in the WT dataset suggested that the inhibitor exhibited high specificity toward WT PLK4, with few off-target effects (as seen in the G95R dataset). As this was a discovery experiment aiming to identify novel PLK4 regulated proteins and phosphosites, where very few substrates are currently known, the decision was made to relax the adj. *p* value to ≤0.075.

DAVID functional annotation analysis of the significantly regulated WT proteins (adj *p* value ≤0.075) revealed an enrichment of proteins at a number of cellular compartments including exosomes, mitochondria and focal adehesions. In addition, dysregulated proteins were involved in biological processes including rRNA processing and negative regulation of apoptosis (Figure 5.8 A). It is likely that a four-hour incubation period is insufficient to allow for an accurate assessment of PLK4 regulation of protein expression due to the varied rates of protein turnover in the cell. However, a number of individual cell cycle-regulated proteins, including cyclin D2, p21 and CDK4 were identified as significantly regulated and are discussed in more detail in section 5.2.5.

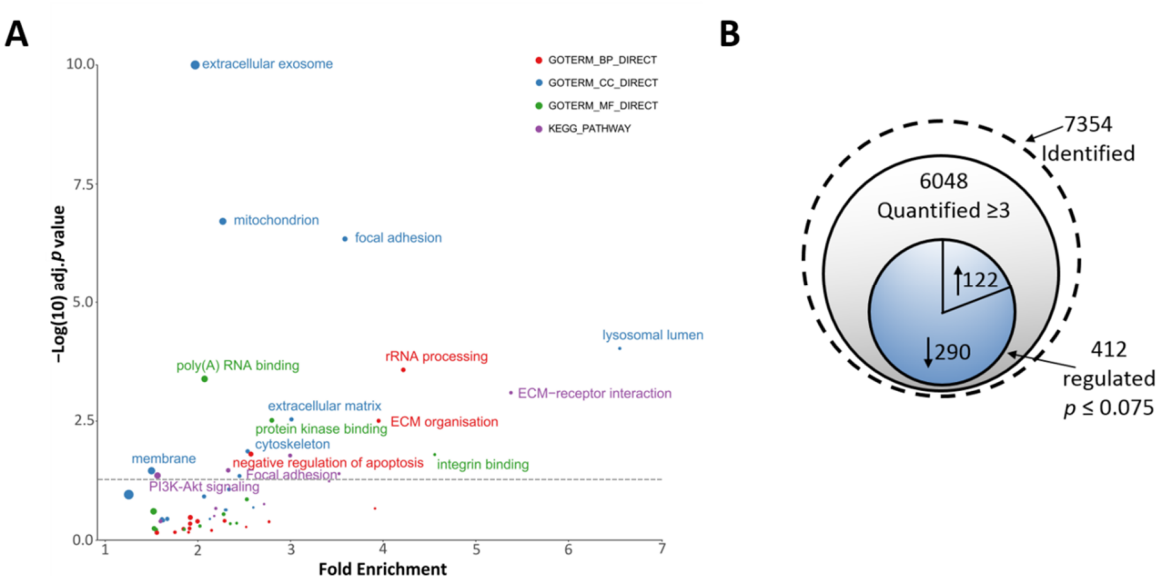


Figure 5.8. Enrichment analysis of significantly regulated proteins in the WT PLK4 dataset. A) GO term enrichment analysis of significantly regulated proteins. Proteins enriched with an FDR adj. *p* value ≤0.05 (using the Benjamini-Hochberg procedure) are shown above the dotted line and labelled. BP= biological processes; CC = cellular compartment; MF= molecular function. B) The total number of proteins identified in the WT PLK4 dataset is shown, along with the total number quantified and proteins identified as significantly regulated ($p \leq 0.075$).

5.2.5. **Assessment of phosphosite regulation**

The work in Chapter 4 described a false localisation rate cut-off to be employed for data sets acquired with various fragmentation modes to ensure high confidence in phosphosite localisation. For HCD OT methods, this cut off was set as 0.994 when phosphosites have been localised using PTM-score. However, as this was a discovery experiment, with the aim of exploring the PLK4-regulated phosphoproteome, a 0.994 cut-off was considered too stringent. Therefore, the widely used PTM-Score of ≥0.75, used to describe highly confident 'class I' phosphosites was employed (Olsen *et al.*, 2006). To exploit the greater capability of ptmRS to confidently localise phosphosites compared with PTM-score, ptmRS scores are also included for the regulated phosphosites discussed in this chapter. The phosphopeptide lists were expanded to assess quantification at the phosphosite level. This is important, as multiply phosphorylated peptides may be differentially regulated and two or more phosphosites within a single peptide may have distinct cellular functions (Lanucara *et al.*, 2016).

Bayesian statistical analysis revealed 140 phosphosites significantly changing with an adjusted *p* value ≤0.05 in the WT dataset (Figure 5.9 A). Similar to the proteomics data, a much greater number of significantly regulated phosphosites were identified in the G95R PLK4 dataset, with 1501 identified with an adj. *p* value ≤0.05 (Figure 5.9 B; Figure 5.10 B).

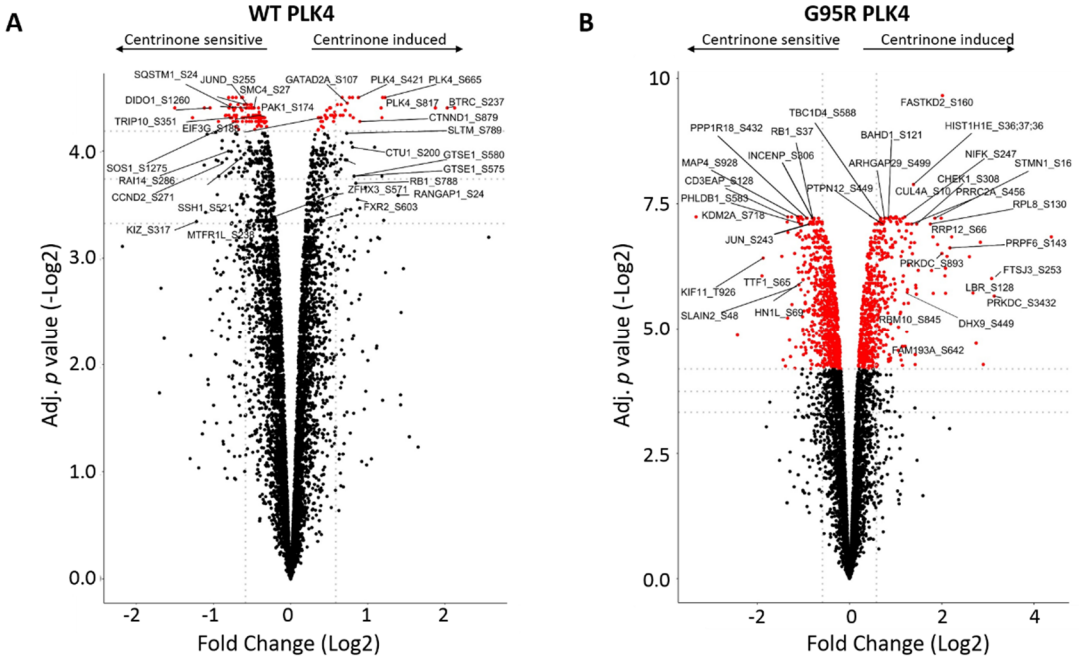


Figure 5.9. Phosphosite analysis of centrinone treated WT- and G95R PLK4 expressing cells. H/L ratios were tested for significance using a Bayesian statistical approach. Log2 fold change of identified phosphosite was plotted against –Log2 Benjamini-Hochberg FDR adj. *p* values for both the WT (A) and G95R (B) PLK4 datasets. Proteins with an adj. p value ≤0.05 are highlighted in red. The horizontal grey lines represent adj. *p* values of ≤0.05 (top), ≤0.075 (middle), and ≤0.1 (bottom).

The phosphosites that changed significantly upon centrinone treatment in the G95R PLK4 expressing cells were found to be enriched in proteins localising to a diverse range of cell compartments, including focal adhesions, the cytoskeleton, kinetochores, nuclear envelope and the midbody of dividing cells. In addition, an enrichment of biological processes, including cytoskeletal organisation, sister chromatid cohesion, spindle organisation and nuclear envelope assembly was observed. This enrichment of late mitotic events corroborates the protein expression data, which revealed significantly regulated proteins with roles in the G2/M phase of the cell cycle (Figure 5.10 A; Figure 5.5 B). The distribution of Ser, Thr and Tyr downregulated phosphosites was compared against the global distribution of identified phosphosites (Figure 5.10 C). As expected, the majority of phosphosites observed were at Ser residues (92.6 %), followed by Thr (7.1 %), with very

few pTyr residues (0.3 %) (Olsen *et al.*, 2006). The regulated phosphosites followed this same trend, with the majority being located on Ser residues (90.9 %).

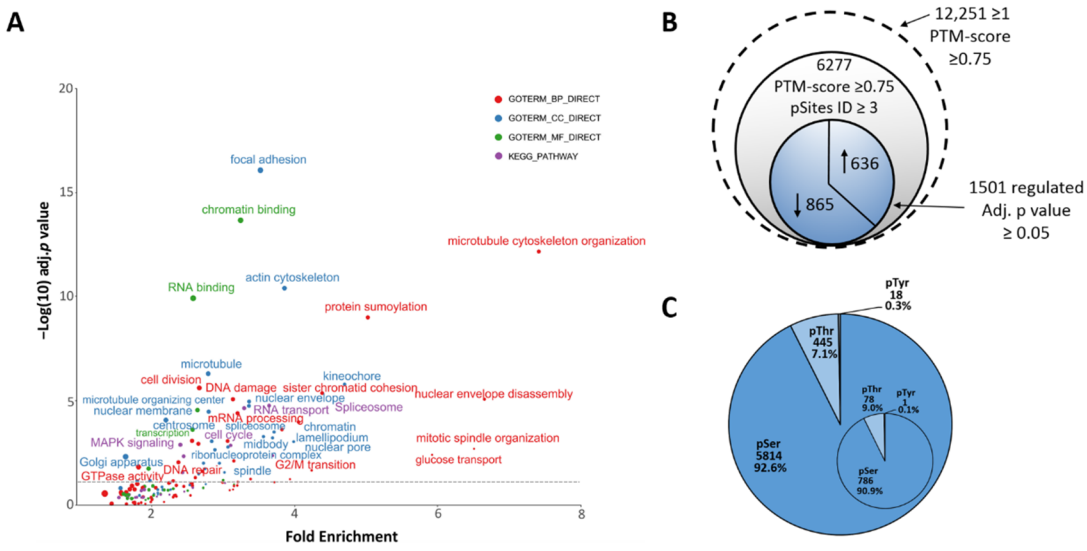


Figure 5.10. Enrichment analysis of significantly regulated phosphosites in the G95R PLK4 dataset. A) GO term enrichment analysis of significantly regulated phosphosites. Phosphoproteins enriched with an FDR adj. *p* value ≤0.05 (using the Benjamini-Hochberg procedure) are shown above the dotted line and labelled. BP= biological processes; CC = cellular compartment; MF= molecular function. B) The total number of phosphosites identified in the G95R PLK4 dataset is shown, along with the total number quantified and phosphosites identified as significantly regulated ($p \leq 0.075$). C) Global distribution of Ser/Thr/Tyr residues and the distribution of significantly downregulated Ser/Thr/Tyr residues (inset).

Taken together, the G95R PLK4 proteomics and phosphoproteomics datasets strongly suggest that in the presence of overexpressed drug-resistant PLK4, centrinone treatment results in a high concentration of unbound inhibitor, which acts off-target, causing widespread changes in the cell. An overlap between the regulated proteins and phosphosites may therefore be identified between the WT and G95R datasets but have occurred via different pathways, with redundancy in the proteins involved. Therefore, the G95R PLK4 dataset cannot be used to validate WT PLK4 substrates and regulated pathways.

5.2.6. **Assessment of WT PLK4 regulated phosphoproteins**

As only ~1.8 % of phosphosites were significantly regulated in the WT dataset and this was a discovery experiment, the adj. *p* value was relaxed to ≤0.075, with the same justification as that used for the proteomics dataset. The adj. *p* value for each phosphosite quantified is listed in Appendix 3 (Table 9.4, 9.5). The distribution of Ser, Thr and Tyr downregulated

phosphosites was compared against the global distribution of identified phosphosites (Figure 5.11 C). The regulated phosphosites followed this same trend as the global distribution, with the majority of regulated phosphosites being located on Ser residues (90.7 %). Functional annotation and enrichment analysis of the significantly regulated phosphoproteins in the WT PLK4 dataset was performed using DAVID. An overview of the enriched terms within the dataset is shown in Figure 5.11 A. Regulated phosphoproteins were identified as localising at the cytoskeleton, chromosomes, and the centrosome. An enrichment of biological processes including cell-cell adhesion, transcription, mRNA processing and regulation of the cell cycle was observed.

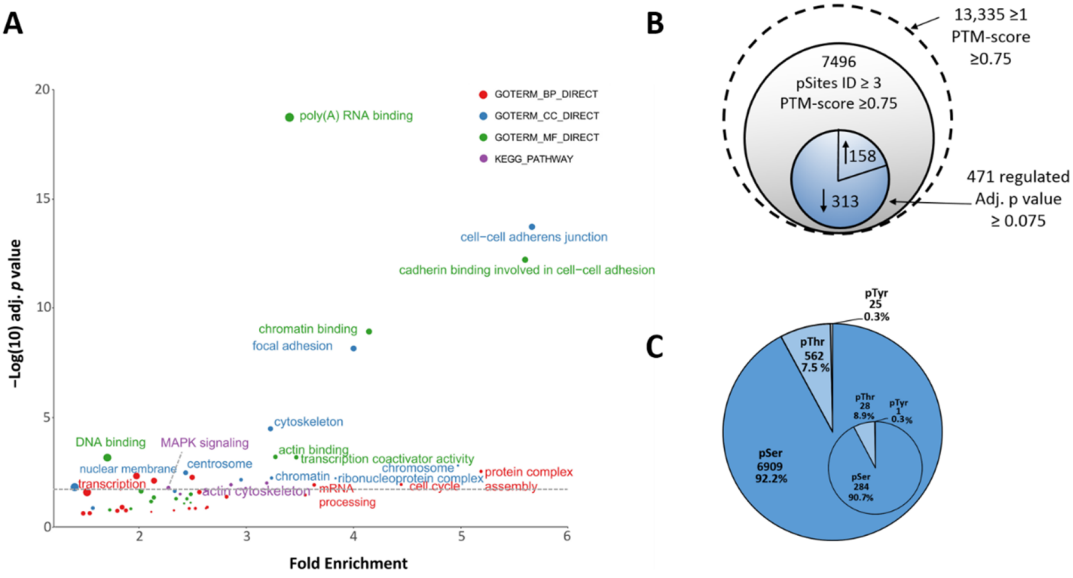


Figure 5. 11. Enrichment analysis of significantly regulated phosphosites in the WT PLK4 dataset. A) GO term enrichment analysis of significantly regulated phosphosites. Phosphoproteins enriched with an FDR adj. *p* value ≤0.05 (using the Benjamini-Hochberg procedure) are shown above the dotted line and labelled. BP= biological processes; CC = cellular compartment; MF= molecular function. B) The total number of proteins identified in the WT PLK4 dataset is shown, along with the total number quantified and proteins identified as significantly regulated ($p \leq 0.075$). C) Global distribution of Ser/Thr/Tyr residues and the distribution of significantly downregulated Ser/Thr/Tyr residues (inset).

The MAPK signalling pathway was identified as misregulated with five members of the pathway identified (Figure 5.12 A, B). Total EGFR protein and pThr669 EGFR were identified as upregulated with centrinone treatment. EGFR Thr669 is located in the juxtamembrane domain and is a substrate of ERK. ERK is activated upon phosphorylation of Tyr204, located within its activation loop and this phosphorylation site was also identified as upregulated in

this study (Butch & Guan, 1996). ERK phosphorylation of pThr669 EGFR results in negative feedback control of EGFR tyrosine kinase activity, preventing autophosphorylation (Nishimura *et al.*, 2009; Sato *et al.*, 2013). An increase in active ERK, along with an increase in pThr669 EGFR suggests that centrinone treatment resulted in increased activity of ERK toward pThr669 EGFR leading to an inhibition of EGFR activity with downstream effects on the MAPK signalling pathway. SOS1 pSer1275 and MEK2 Ser23 also identified in the MAPK pathway have only been identified previously in high-throughput phosphoproteomics with their biological function not yet known. This work implicates PLK4 inhibition in regulation of the MAPK pathway, which is directly linked to regulation of the cell cycle (Wilkinson & Millar, 2000)

Regulated phosphosites and proteins were manually mapped on to the cell cycle, which revealed proteins and phosphoproteins implicated in G1 and S phase (Figure 5.12 A, B). This included downregulated cyclin D2, CDK4, CDK6 and p21 proteins (adj. *p* value ≤0.075). The activity of cyclin D2- CDK4 complexes are suppressed by p21, which prevents phosphorylation of RB1 required for G1/S progression (Xiong *et al.*, 1993). Paradoxically, p21 can also act as an adaptor of cyclin D-CDK4 complexes, assembling the active complex in the nucleus to promote cell proliferation (LaBaer *et al.*, 1997). The lower expression of these proteins upon centrinone treatment is likely to exert an anti-proliferative effect on the cells, mediated by PLK4 inhibition.

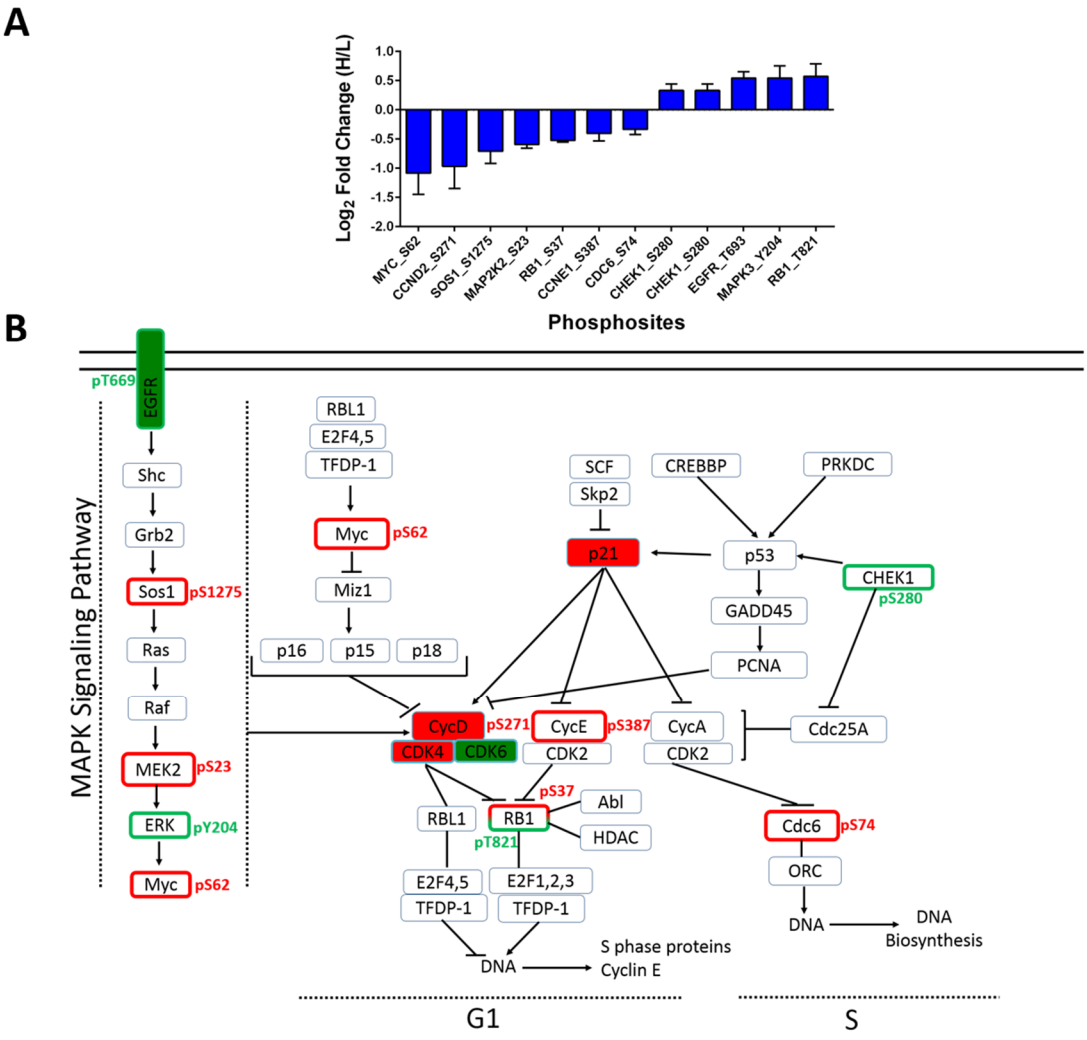


Figure 5.12. Pathway analysis of significantly regulated phosphosites in the WT PLK4 dataset. A) Bar chart representing significantly regulated (adj. *p* value ≤0.075) phosphoproteins implicated in MAPK and cell cycle pathways. Bar charts represent the mean Log2 fold change (H/L) for four bioreplicates ±SD. B) significantly regulated (adj. *p* value ≤0.075) proteins/phosphoproteins were identified as components of the MAPK signalling pathway. Additionally, significantly regulated proteins/phosphoproteins (adj. *p* value ≤0.075) were manually mapped on to the cell cycle and were implicated in G1/S phase. Green outline= phosphosite upregulated; red outline= phosphosite downregulated; block green= protein upregulated; block red= protein downregulated.

Interestingly, two RB1 phosphosites were identified as differentially regulated, with a decrease in pSer37 and increase in pThr821 observed. RB1 protein was identified with no significant change in expression and therefore, the two identified phosphosites reflect differential regulation of phosphorylation. Whilst pSer37 has only been previously identified in high-throughput phosphoproteomics studies, pThr821 has established roles in RB1 activity. RB1 binds to proteins with a specific LxCxE motif to prevent cell growth and this binding is dependent on dephosphorylation of Thr821. Phosphorylation of Thr821 inhibits RB1 binding to proteins and therefore, the increase in phosphorylation observed in this study suggests a positive effect on cell proliferation (Knudsen & Wang, 1997).

Cdc6 Ser74 was significantly downregulated 1.3 fold with centrinone treatment. Cdc6 is a substrate of PLK4 and has a role in negatively regulating the PLK4-STIL-SAS6 complex required for centriole duplication. During G2 phase of the cell cycle, Cdc6 is recruited to the centrosomes and binds to SAS6, blocking its interaction with STIL and preventing centrosome overduplication. This interaction is disrupted in S phase by PLK4 phosphorylation of Cdc6 at two sites, pSer30 in its N-terminal domain and pThr527 (not identified in this study). In this way, PLK4 and Cdc6 antagonistically control centrosome amplification to ensure the centrioles are duplicated once per cycle (Xu *et al.*, 2017). The observation of downregulated pSer74 Cdc6 suggests that this could be a novel target of PLK4 within the N-terminal domain.

5.2.7. **Identification of PLK4 phosphorylation sites**

Three PLK4 phosphosites, Ser421, Ser665 and Ser817 were identified as upregulated upon treatment with centrinone, with H/L ratios of 1.8, 2.3 and 3.9 respectively (Figure 5.13). For pSer421 and pSer665, the H/L ratios match the change observed in protein expression, suggesting that there was no stoichiometric increase in phosphorylation at these sites. A stoichiometric increase is observed for pSer817, suggesting that this phosphosite is targeted by an upstream protein kinase. The identified phosphosites are located in the C-terminus, with Ser421 located in the PB1 domain and Ser665 located in PB3. Ser817 is located at the extreme C-terminus, outside of PB3. While all three sites have been previously observed in high-throughput proteomic studies, no functional role of phosphorylation at these sites is currently known.

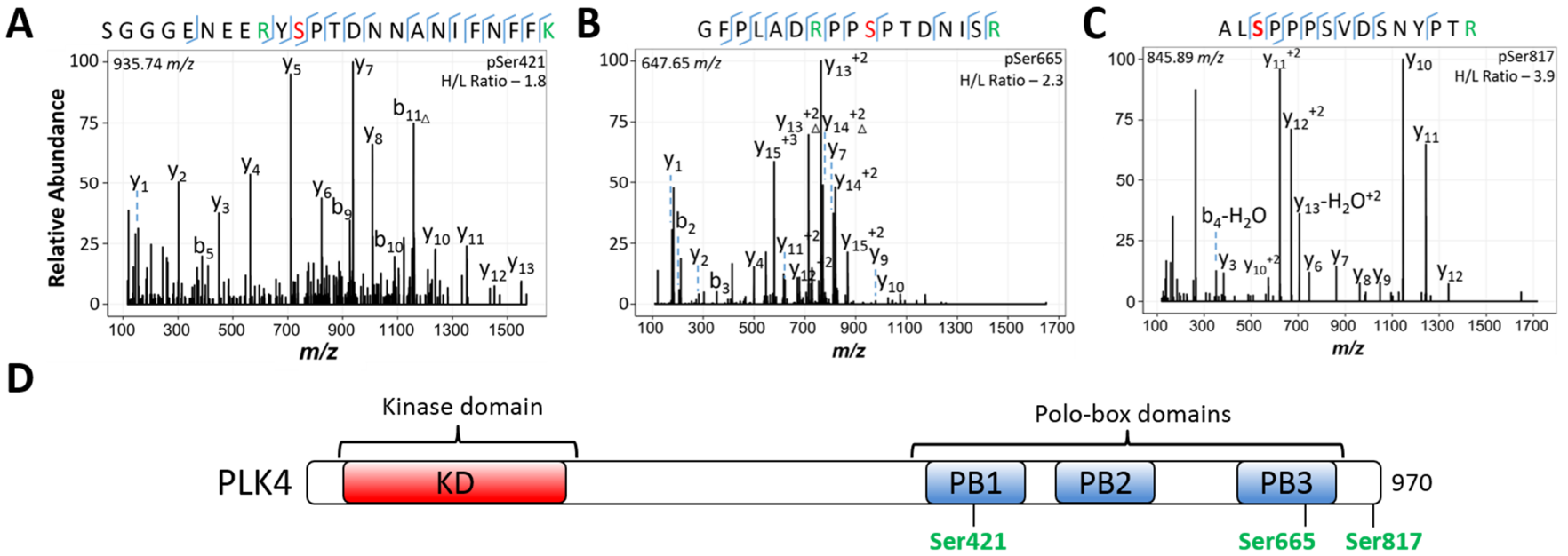


Figure 5.13. Identification WT PLK4 phosphosites *in vitro*. Three significantly upregulated (adj. *p* value ≤0.075) PLK4 phosphosites were identified. The peptide sequence and the identified phosphosite is detailed on each mass spectrum. A) Doubly charged ion at *m/z* 935.74, identifying pSer421. B) Doubly charged ion at *m/z* 647.65, identifying pSer665. C) Doubly charged ion at 845.89, identifying pSer817. D) PLK4 schematic highlighting the location of the identified phosphosites.

5.2.8. **Identification of previously described PLK4 substrates & interactors**

To date, only a limited number of PLK4 substrates have been identified. Protein substrates STIL and PCM1 were identified in the total protein dataset. However, neither were identified as significantly regulated following centrinone treatment of WT PLK4 expressing cells (Table 5.1). Cep192, a key PLK4 scaffold protein was identified as significantly downregulated, with an adj. *p* value of ≤0.072. Phosphorylated forms of CP110, PCM1, Cep152/Cep192 and FBXW5 were identified however, none of the identified phosphosites were significantly regulated and the phosphosites that were identified are not known to be phosphorylated by PLK4.

Table 5.1. Identification of previously described PLK4 substrates. Known PLK4 substrates are shown, that were identified at the protein or phosphosites level. Associated adj. *p* values and H/L ratios are listed. For phosphosites ID's, PTM-score and ptmRS site localisation scores are also shown.

	Protein ID	pPeptide ID	PTM-score	ptmRS score	Adj *p* value	H/L ratio
STIL	Yes	No	NA	NA	0.746	1.04
CP110	No	S641	1.000	100	0.849	0.97
PCM1	Yes	NA	NA	NA	0.277	0.97
		S110	1.000	100	0.877	1.02
		S960	1.000	100	0.840	1.02
		S1257	1.000	100	0.924	0.66
		S1730	1.000	100	0.668	0.94
		S1765	1.000	100	0.613	1.06
		S1768	1.000	99.99	0.613	1.06
		S1958	1.000	100	0.965	1.01
GCP6	Yes	No	NA	NA	0.166	0.94
Cep152	No	S1461	1.000	100	0.111	0.86
Cep192	Yes	NA	NA	NA	0.072	0.92
		S1159	1.000	100	0.751	1.02
FBXW5	No	S275	0.935	100	0.922	1.02
		S284	1.000	100	0.131	1.46

In addition, the BioGRID database was searched for known PLK4 interactors and cross-referenced against the list of significantly regulated proteins and phosphoproteins. A number of PLK4 interactors were identified as significantly changing, including BTRC, MAGED1, NEDD1 and phosphorylated MYCBP2 (Ser3467) (Table 5.2) (Firat-Karalar *et al.*, 2014).

Table 5.2. Identification of previously described PLK4 interactors. Known PLK4 interactors in the BioGrid database are shown if they were identified as significantly regulated at the protein or phosphosites level (adj. *p* values ≤0.075). Adj. *p* values, H/L ratios and associated PTM-score and ptmRS site localisation scores are listed.

	Protein ID	pPeptide ID	PTM-score	ptmRS score	Adj. *p* value	H/L ratio
MYCBP2	No	S3467	1.000	100	0.063	1.26
BTRC	Yes	NA	NA	NA	0.065	0.84
		S237	0.998	NA	0.047	4.08
MAGED1	Yes	No	NA	NA	0.021	0.81
NEDD1	Yes	No	NA	NA	0.063	0.92

NEDD1 is a component of the PCM and binds to CEP192. At the centrosome, NEDD1 recruits γ-tubulin to nucleate the microtubules in interphase (Manning *et al.*, 2010). It is therefore possible that PLK4 interacts with NEDD1 via their interaction with the Cep192 scaffold. MYCBP2 is an atypical RING E3 ubiquitin ligase that adds ubiquitin groups to Thr rather than Lys residues (Pao *et al.*, 2018) and has a role in cell growth via regulation of mTOR pathway through the ubiquitylation of target proteins (Han *et al.*, 2012). MAGE proteins can assemble with RING ubiquitin ligases to regulate ligase activity and alter subcellular localisation (Lee & Potts, 2017). The identification of MYCBP1 and MAGE as significantly upregulated in this study suggests they are not direct PLK4 substrates but do have potential roles at the centrosome that may be regulated by PLK4 activity.

BTRC (also known as β-TrCP), the substrate recognition component of the SCF E3 ubiquitin ligase complex required for PLK4 degradation was identified as significantly regulated at both the protein and phosphosite level (Guderian *et al.*, 2010). Protein expression was downregulated upon centrinone treatment; with a 1.2 fold decrease observed. At the phosphosite level, pSer237 of β-TrCP was identified as a novel phosphosite, with a 4-fold increase upon centrinone treatment, reflecting a stoichiometric increase in phosphorylation (Table 5.2). Inhibition of PLK4 activity with centrinone prevents autophosphorylation within the degradation motif and protects the protein from β-TrCP mediated proteasomal degradation. Therefore, the reduced expression of β-TrCP expression when PLK4 activity is inhibited is expected. The identification of increasing levels of a novel phosphosite is unexpected and requires further investigation to elucidate its biological function.

5.2.9. **Enrichment of PLK4 regulated phosphoproteins at the centrosome**

As the major function of PLK4 is at the centrosome to co-ordinate centriole biogenesis, it was not surprising that centrosome localisation was revealed as an enriched term in Figure 5.11 A. In total, 22 significantly regulated phosphoproteins were identified as enriched at the centrosome, with 14 downregulated and 8 upregulated (Figure 5.14 A). The full list of centrosomal regulated phosphoproteins and their associated adj. *p* value can be found in Appendix 2.

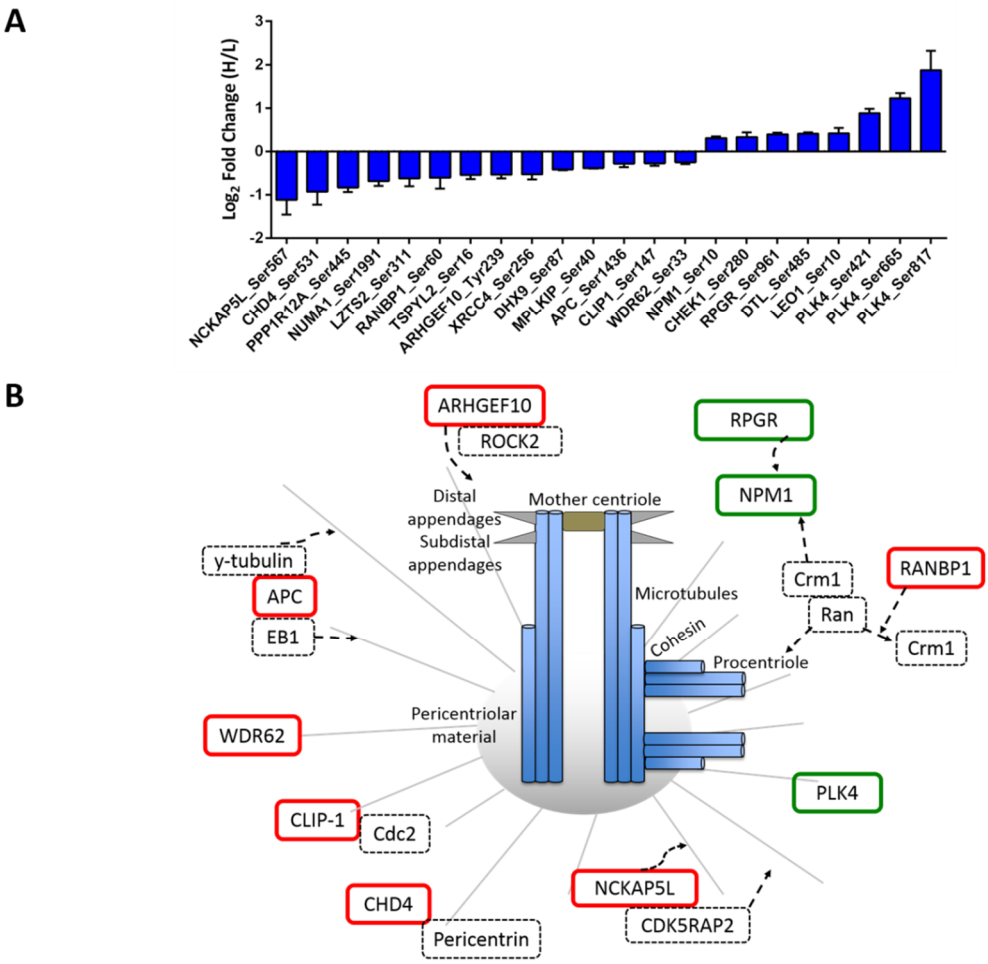


Figure 5.14. Identification of centrosomal-associated phosphoproteins. A) Analysis of significantly regulated (adj. *p* value ≤0.075) phosphosites in DAVID revealed an enrichment of phosphoproteins localised at the centrosome. Bar charts represent the mean Log2 fold change (H/L) for four bioreplicates ±SD. B) Schematic to highlight the roles of the identified phosphoproteins in centrosome biology. Red outline= downregulated with centrinone treatment; green outline= upregulated with centrinone treatment.

NCKAP5L (also known as Cep169) pSer567 was identified as significantly downregulated following treatment of the cells with centrinone (Figure 5.14 A, B). Cep169 binds to the microtubule plus-end tracking protein, CDK5RAP2, which is responsible for recruiting γ-TuRC complexes to the centrosome to nucleate growing microtubules (Mori *et al.*, 2015a). Phosphorylation at Ser567 has only been previously identified in high-throughput phosphoproteomics studies, although this work suggests that this is potentially regulated by PLK4. Adenomatous polyposis coli (APC), a tumour suppressor, which localizes to the centrosomes, was downregulated at Ser1436. APC functions at interphase centrosomes and interacts with γ-tubulin to stimulate growth of the nascent microtubules (Lui *et al.*, 2016). Whilst a functional understanding of the role Ser1436 has not yet been defined, downregulated APC phosphorylation at the centrosome suggests that this is also a possible substrate of PLK4.

Phosphorylated CHD4 (pSer531), a component of the multiprotein nucleosome remodelling deacetylase (NuRD) complex, was also identified as downregulated (Figure 5.14 A, B). CHD4 localises at centrosomes and has been shown to negatively regulate the centrosomal localisation of pericentrin, a key component of the PCM responsible for anchoring proteins, including γ-TuRC complexes to the centrosome (Doxsey *et al.*, 1994) (Sillibourne *et al.*, 2007). Ser531 phosphorylation is frequently observed in high-throughput phosphoproteomics studies, although no functional analysis has been undertaken. Downregulated phosphorylation of Tyr239 in Rho guanine nucleotide exchange factor ARHGEF10 was also identified in this study. ARHGEF10 localises at the centrosome during G1/S and regulates RhoA signalling which is believed to have a role in centriole duplication (Aoki *et al.*, 2009).

Phosphorylated NPM1 (pSer10) was identified as significantly upregulated, although the protein level was also significantly upregulated, which suggests that there was no stoichiometric increase in the level of phosphorylation. NPM1 has established roles at the centrosome in negatively regulating centriole duplication (Okuda *et al.*, 2000). NPM1 is a substrate of the Ran-Crm1 network, which regulates NPM1 centrosomal localisation via binding at its nuclear export sequence to prevent unscheduled centriole duplication (Wang *et al.*, 2005). Phosphorylated RANBP1 (Ser60) was also identified in this study although no functional role has been attributed to this phosphosite. RANBP1 is also linked to the Ran-

crm1 complex as it dissociates Crm1 from Ran which prevents the centrosomal localisation of NPM1, leading to unscheduled centriole duplication (Budhu & Wang, 2005).

NPM1 also interacts with RPGR, a possible guanine-nucleotide release factor, which is suggested to be involved in dissociating NPM1 from the centrioles (Shu *et al.*, 2005). RPGR phosphorylation at Ser961 was identified as significantly upregulated and therefore unlikely to be a direct substrate of PLK4. Similarly, NPM1 is unlikely to be a direct substrate of PLK4, as pSer10 is upregulated following PLK4 inhibition. As NPM1 is a negative regulator of centriole duplication, it may be upregulated to maintain suppression of centriole biogenesis in the absence of active PLK4.

5.2.10. **PLK4 consensus motif analysis**

Next, the motifs surrounding the significantly downregulated phosphosites were examined to assess whether a PLK4 phosphorylation consensus sequence could be identified. The seven residues either side of the phosphosite were searched using the iceLogo tool (Colaert *et al.*, 2009). This analysis revealed a motif enriched for proline residues. This included the enrichment of Pro at the +1 position.

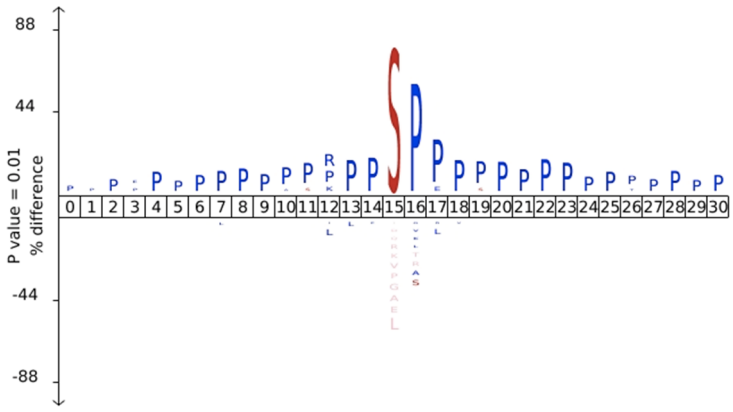


Figure 5.15. Identification of novel PLK4 phosphorylation motifs. Downregulated PLK4 phosphosites were analysed using iceLogo to identify phosphorylation consensus motifs. This analysis identified a proline directed phosphorylation motif (p = 0.01).

The autophosphorylation analysis of PLK4 catalytic domain presented in Chapter 3 revealed autophosphorylation at Ser232, which contains a Pro at the -1 position. In addition, the FLAKSFG**S**PNRAYKK peptide substrate (P32) used as our standard PLK4 substrate for *in vitro*

kinase assays (Chapter 3) also already includes a Pro at the +1 relative to the phosphorylation site, providing further evidence that PLK4 phosphorylates proline-directed motifs. To further evaluate PLK4 phosphorylation of sites similar to the enriched motif identified as centrinone-sensitive in cells, the peptide substrate termed P32 (see Chapter 3, Figure 5.16 A) was modified at a number of positions, and the phosphotransferase activity of recombinant PLK4 assessed using them as substrates (Figure 5.16 A). The rate of phosphorylation by PLK4 towards the +1 NT (+1 meaning 1 extra Pro at the N-terminus) peptide was essentially identical to that of the original SP motif-containing peptide (Figure 5.16 B), confirming that the PXP motif is phosphorylated by PLK4 *in vitro*. Importantly, PLK4 exhibited very little activity toward the +1 CT (+1 meaning 1 extra Pro at the C-terminus) peptide. This confirmed that the +1 proline is critical for PLK4 phosphorylation of this peptide sequence. Strikingly, if an additional proline was incorporated at the -2 position to generate a PPSP motif (+2 NT), PLK4 activity was completely abolished (Figure 5.16 B), and this sequence motif was also excluded in cellular substrates sensitive to centrinone. An ATP titration experiment was performed to determine *K*m values for the P32 peptide and the +1NT peptide. The *K*m[ATP] values for peptide 32 (GXP) and +1NT (PXP) were calculated to be 12.8 µM and 10.8 µM respectively (Figure 5.16 C).

As Aurora A is also inhibited by centrinone *in vitro* (Chapter 3) an additional assay was performed to measure whether the proline-directed motif in P32 peptide was phosphorylated by Aurora A. Figure 5.16 D shows that Aurora A has extremely low activity against the 5-FAM-FLAKSFG**S**PNRAYKK peptide, and no detectable activity toward the +1NT peptide. This provides further evidence that PLK4 activity is directed toward proline motifs and suggests that the downregulated cellular phosphosites observed in this study are due to inhibited PLK4 (rather than Aurora A) activity. Taken together, the *in vitro* kinase assays provide strong evidence PLK4 phosphorylates a proline-directed consensus motif, which will be useful in the future for trying to understand PLK4 substrates in cells.

A

Name	Sequence
P32	5-FAM-FLAKSFGSPNRAYKK-Amide
+1 NT Pro	5-FAM-FLAKSFPSPNRAYKK-Amide
-1 CT Pro	5-FAM-FLAKSFPSANRAYKK-Amide
+2 NT Pro	5-FAM-FLAKSPPSPNRAYKK-Amide

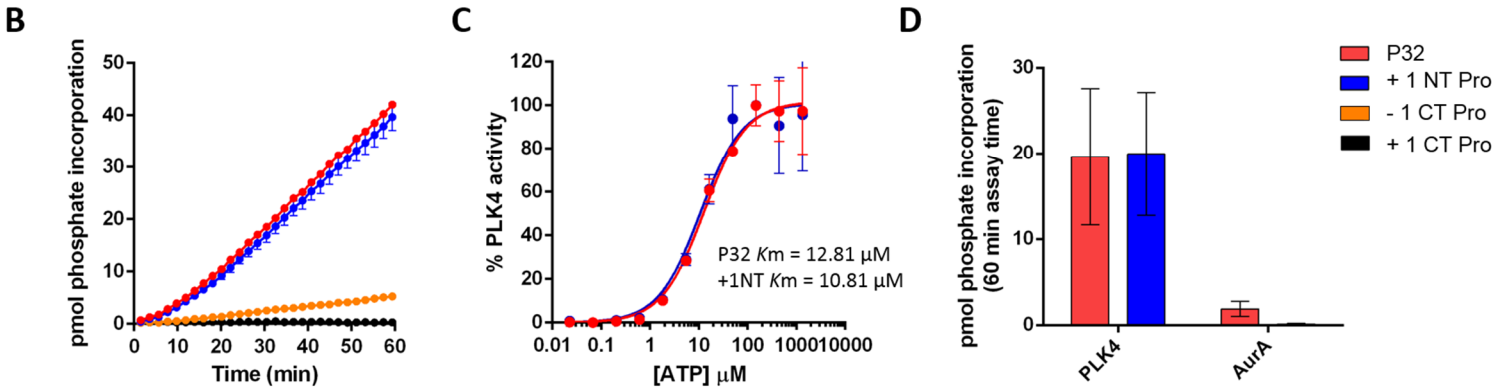


Figure 5.16. Validation of PLK4 phosphorylation consensus motif. A) The commercially available P32 peptide substrate was used to design peptides incorporating or removing additional prolines residues to test the phosphorylation motif. NT= N-terminal to the phosphosites; CT= C-terminal to the phosphosites. B) WT PLK4 (1-269) was assayed by quantified mobility shift using EZ Reader technology with a fluorescent peptide substrate in the presence of 1 mM ATP (final concentration), to mimic cellular ATP levels. C) *K*m of WT PLK4 (1-269) was determined for both P32 and the +1 NT following titration of increasing concentrations of ATP. D) Recombinant WT Aurora A and PLK4 (1-269) were assayed to test Aurora A activity toward the P32 and +1 NT Pro peptides in the presence of 1 mM ATP. The activity assays shown were performed by Dr. Dominic Byrne at the University of Liverpool.

5.2.11. **Validation of selected regulated proteins & phosphosites**

Validation of selected PLK4 regulated proteins and phosphosites was performed by western blot analysis. Cyclin D2 and p21 expression was analysed to assess whether downregulation following centrinone treatment could be observed by this method. Figure 5.17 shows that for cyclin D2 (CCND2), a decrease in band intensity is observed when the cells are incubated with centrinone. The data becomes less clear however, as a decrease is also observed with centrinone treatment without induction of expression of PLK4. This could be due to off-target effects also resulting in down regulation of cyclin D2. Alternatively, the 'leaky' expression of PLK4 (shown previously in Chapter 3) may be sufficient to induce an effect. Centrinone does not appear to have any effect on the levels of CCND2 observed with or without G95R PLK4 expression. If decreases in CCND2 were due to off-target effects then this should also be observed in the G95R PLK4 U2OS cells with and without induction of protein expression. However, as no decrease in CCND2 levels can be observed, it is likely that the downregulation observed is directly linked to the inhibited activity of PLK4. No decrease in the intensity of p21 can be observed in WT or G95R PLK4 expressing cells treated with or without centrinone. This is unexpected; p21 was identified as significantly downregulated, with a 1.8 fold decrease in expression in the proteomics dataset. It is possible that this fold change is too small to be observed by western blot. This experiment would need to be repeated to validate p21 as a downstream target of PLK4 activity and may require longer incubation times with centrinone to reduce p21 levels, such that a decrease can be observed by western blot.

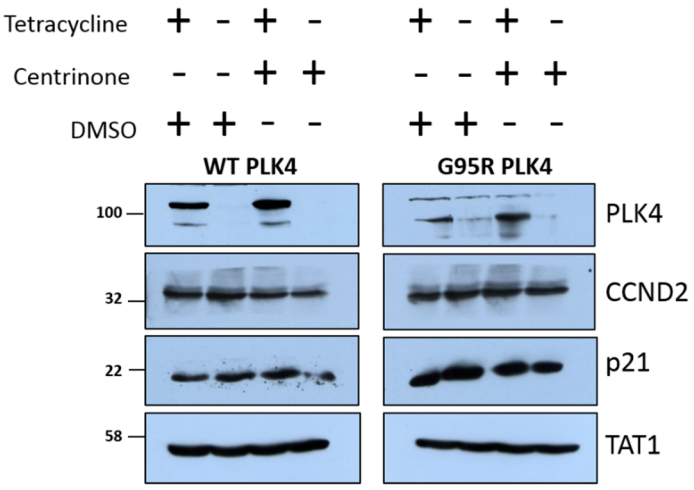


Figure 5.17. Cellular analysis of potential downstream PLK4 substrates. Expression of WT & G95R PLK4 was induced upon incubation with 1 µg/mL tetracycline for 18 hours. Cells were incubated with 300 nM centrinone, or 0.1 % DMSO for four hours. Total lysates were analysed by western blot using the indicated antibodies.

Next, upregulation of total EGFR and pThr669 was evaluated by western blotting, using total and phospho-specific EGFR antibodies (Figure 5.18). As the fold change observed in the MS data was reasonably small (1.2 fold increase for total EGFR and 1.5 fold for pThr669), cells were induced to express PLK4 and treated with centrinone for 4, 16 and 24 hours in an attempt to amplify the effect of centrinone treatment. No changes in total EGFR band intensity could be observed, which may indicate that the fold change is too small to observe by western blot, or that EGFR is regulated only at the level of phosphorylation.

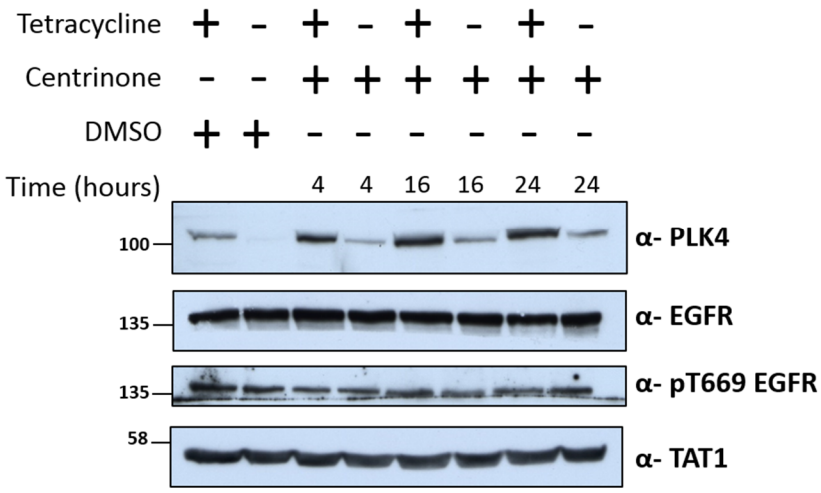


Figure 5.18. Analysis of potential PLK4 substrates in U2OS stable cells. Expression of WT PLK4 was induced upon incubation with 1 µg/mL tetracycline for 18 hours. Cells were incubated with 300 nM centrinone, or 0.1 % DMSO for four hours. Total lysates were analysed by western blot using the indicated antibodies.

A decrease in pThr669 band intensity was observed following incubation of cells with centrinone for four hours (lanes 3 and 4) compared with the untreated cells (Figure 5.18). However, the decrease was observed irrespective of PLK4 overexpression, suggesting this is an off-target effect of centrinone inhibition. At 16 hours (lanes 5 and 6), an increase in pThr669 is observed compared to inhibition at 4 hours, although the levels are slightly lower than those observed in the untreated cells (lanes 1 and 2). At 24 hours (lanes 7 and 8), the levels of Thr669 phosphorylation observed in WT PLK4 overexpressing cells treated with centrinone is again lower than those observed in the untreated cells, although an increase in phosphorylation is observed at 24 hours in non-PLK4 overexpressed cells treated with centrinone. Taken together, the data highlights the true complexity of cell signalling mechanisms and illustrates the likelihood of redundancy in the number of protein kinases capable of phosphorylating and regulating this residue in EGFR.

5.3. Conclusions

The aim of this work was to identify novel PLK4 substrates and regulated signalling pathways to provide a deeper understanding of the role of PLK4 in co-ordinating cellular activities. Utilising both total proteomics and phosphoproteomics led to the identification of a number of significantly changing proteins and phosphosites, implicating PLK4 activity in a range of biological functions, including those beyond centrosome biogenesis.

Treatment of G95R PLK4 expressing cells with centrinone led to the identification of a much greater number of regulated proteins and phosphosites compared with the WT dataset, with the possibility that the cells had begun to initiate mitotic arrest. The differential effects of centrinone on the two cell lines may be explained by the inability of drug-resistant G95R PLK4 mutant to bind centrinone, leading to an increase of unbound drug (300 nM) in the cell which could then potently inhibit off-target proteins. This includes Aurora A, which in Chapter 3 was shown to be inhibited by centrinone. Inhibition of Aurora A by small molecule inhibitors such as VX-680 and MLN8237 have been shown to lead to mitotic arrest (Gorgun *et al.*, 2010; Li *et al.*, 2010). This offers the possibility that the effects observed in the G95R PLK4 cell line arise from Aurora A inhibition, leading to negative effects on cell proliferation. Reduced expression of mitotic proteins, including cyclin B upon G2/M arrest has been reported previously, which matches the observation of downregulated mitotic proteins in the G95R PLK4 dataset (Singh *et al.*, 2004). It is therefore possible that off-target

inhibition by centrinone has resulted in the onset of G2/M phase arrest in the G95R PLK4 cell line through off-target inhibition of Aurora A.

This chapter described a number of PLK4 regulated proteins with established roles in the early phases of the cell cycle, including RB1, cyclin D, CDK4 and p21. PLK4 expression has been implicated in the regulation of expression of these G1/S phase proteins at the level of transcription. Deficiencies in PLK4 expression in a PLK4 +/- mouse model, resulted in the delayed activation of cyclin D and CDK4, and suppressed the expression of p21 (Ko *et al.*, 2005). This led to delays in cell cycle progression, spindle defects and an increase in hepatocellular carcinomas. The work described in this chapter provides further evidence that PLK4 regulates the expression of these interphase proteins and highlights the importance of studying the wider roles of PLK4 in the cell cycle.

One interesting observation was the identification of significant differential regulation of the two RB1 phosphosites. Particularly, upregulated pThr821 suggests that RB1 is exerting a positive effect on cell growth, as dephosphorylation of this residue is required for protein binding downstream to suppress cell cycle progression. The majority of other G1/S phase proteins and phosphosites were downregulated to suggest a negative effect on cell growth, and this conflict highlights the complexity of cell signalling and its effect on cell cycle regulation. Although RB1 pThr821 dephosphorylation is required to bind a subset of proteins to exert an inhibitory effect, it is inconsequential for binding to E2F proteins, which are the major downstream binding partners of RB1 (Knudsen & Wang, 1997). RB1 is phosphorylated at a number of residues, with different combinations mediating RB1 activity in a variety of cellular functions (Rubin, 2013). Therefore, the identification of upregulated pThr821 RB1 alone may not be sufficient to confer a positive effect on cell growth. Overall, the trend toward downregulated G1/S phase proteins/phosphosites following WT PLK4 inhibition with centrinone suggests that inhibited PLK4 activity leads to a negative effect on cell growth during G1 phase.

This work also revealed three potential PLK4 phosphosites, all located at the regulatory C-terminal domain. Although the phosphosites have been identified in shotgun phosphoproteomics studies previously (Hornbeck *et al.*, 2015), no function has yet been attributed to them. PLK4 Polo Box 1 (PB1) has been implicated in binding to target proteins and, along with PB2, is necessary for correct localisation at the centrioles and for forming PLK4 homodimers (Slevin *et al.*, 2012). Therefore, it is possible that pSer421, located in PB1

is involved in these functions. PB3 has been shown to relieve linker 1 (L1) mediated PLK4 auto inhibition (Klebba *et al.*, 2015). In addition, it is required for binding to STIL to initiate centriole biogenesis (Ohta *et al.*, 2014). Therefore, the identification of pSer665 located within PB3 may be implicated in either of these functions.

Finally, this work revealed a potential novel PLK4 phosphorylation consensus motif. The currently accepted PLK4 consensus was derived through screening PLK4 against a library of peptides derived from known human phosphorylated peptides. This determined that the PLK4 consensus sequence included as basic residue at -3, an acidic residue at -2, Tyr or a hydrophobic residue at +1 and +2 and Ser or Thr at +4 (Leung *et al.*, 2002; Johnson *et al.*, 2007). Interestingly, Leung et al., showed that proline at the +1 position was not tolerated. However, the data in this thesis provides evidence to support PLK4 phosphorylation of proline-directed motifs both *in vitro* and in cells. The data in Figure 5.16 shows the importance of the +1 proline for PLK4 activity and may support the identified centrosomal proteins being direct targets of PLK4.

Overall, the work in this chapter describes the first large-scale phosphoproteomics study to investigate the global activities of PLK4. The data described reveal potential roles for PLK4 in the regulation of a wider range of centrosomal proteins than previously reported. In addition, the data supports a role of PLK4 in regulating the early stages of the cell cycle. Beyond this, PLK4 activity was implicated in a number of additional biological processes including gene expression and DNA repair. Taken together, this work highlights important novel aspects of PLK4 activity and demonstrates the utility of large-scale phosphoproteomics to investigate the global roles of PLK4, enhancing the understanding of this critical protein kinase.

6. Phosphorylation and activation of Aurora A by PLK4

6.1. Introduction

The previous chapter discussed a large-scale analysis of the PLK4 regulated phosphoproteome, which identified a number of up- and down-regulated proteins and phosphosites with roles in the cell cycle. However, not all previously identified PLK4 substrates could be identified with this approach due to difficulties in identifying low abundance proteins and transient phosphorylation events by shotgun phosphoproteomics (Solari *et al.*, 2015).

Recent work has identified a potential interaction between PLK4 and Aurora A in the formation of acentriolar spindles in mammalian oocytes (Bury *et al.*, 2017). Inhibition of both kinases, with the inhibitor VX-680, resulted in loss of microtubule nucleation, whereas expression of a drug-resistant Aurora A mutant (G216L) restored the spindle structures. Expression of G95R PLK4 failed to rescue the phenotype, however, expression of both drug-resistant mutants together enhanced the rescue of spindles beyond that of G216L Aurora A alone. This strongly suggested a combinatorial role of the activities of PLK4 and Aurora A, although there was no evidence to confirm that the two kinases directly interacted with each other to co-ordinate spindle formation.

In somatic cells, no direct interaction between Aurora A and PLK4 has been previously defined. However, it has been reported that Aurora A directly interacts with another polo-like kinase family member, PLK1 to co-ordinate activities at the centrosome (Asteriti *et al.*, 2015). This highlights the possibility that Aurora A might interact with other members of the family, including PLK4.

As discussed in Chapter 1, Aurora A activity is controlled by phosphorylation at Thr288 within its T-loop, although the mode of activation is complex. Binding partners TPX2, Bora and Ajuba act to induce specific conformational changes that enhance the activity of distinct pools of Aurora A at varying stages of the cell cycle (Eyers *et al.*, 2003; Hirota *et al.*, 2003; Hutterer *et al.*, 2006; Bai *et al.*, 2014). Active Aurora A plays a number of major roles in the cell cycle, including mitotic entry, centrosome maturation, and chromosome

segregation (Giet & Prigent, 1999) with many of these activities requiring interplay with the polo-like kinase family (Asteriti *et al.*, 2015).

Aurora A and PLK1 co-localise during G2 phase and directly interact via Aurora A phosphorylation of PLK1 at Thr210, within the T-loop, leading to activation of PLK1 and mitotic entry (Macurek *et al.*, 2008). This process is greatly enhanced by the Aurora A cofactor, Bora, which interacts with PLK1 to make Thr210 more accessible (Figure 6.1). PLK1 then directs the degradation of Bora via a feedback mechanism dependent on phosphorylation of a DSGxxT motif, promoting binding of the E3 ubiquitin ligase SCF-βTrCP (Seki *et al.*, 2008).

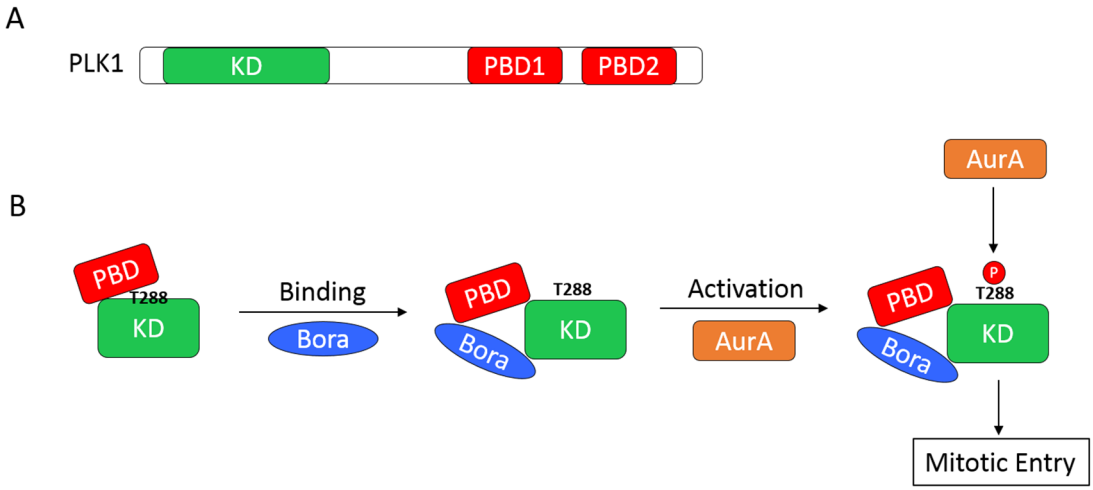


Figure 6.1. PLK1 regulation by Aurora A. A) Schematic of PLK1 showing the kinase domain (KD) & polo-box domains (PBD1- residues 417-480 & PBD2- residues 515 to 584). B) Schematic showing the activation of PLK1 by Aurora A in the presence of Bora. Adapted from the figure in Seki *et al*., 2010.

Once Bora levels have decreased, PLK1 and Aurora A co-localise with the scaffold protein Cep192, bringing both proteins in close proximity (Joukov *et al.*, 2010). This also permits phosphorylation of PLK1 by Aurora A, with an additional role in centrosome maturation through recruitment of γ–tubulin to the centrosome to nucleate the microtubules (Gomez-Ferreria *et al.*, 2007; Joukov *et al.*, 2010; Joukov *et al.*, 2014).

PLK4 is also recruited to Cep192, in addition to another scaffold protein, Cep152. An alpha helical region, present in both Cep152 and Cep192, ensures the specificity of binding to PLK4 whilst a neighbouring motif consisting of acidic residues results in electrostatic

interactions to anchor PLK4 to these proteins (Sonnen *et al.*, 2013). The interaction of both PLK4 and Aurora A with Cep192, and their key roles at the centrosome, suggests that these two kinases could also interact with each other.

No Aurora A phosphopeptides were identified as significantly regulated in the PLK4 phosphoproteomics study in Chapter 5 to support the possibility of PLK4 regulation of Aurora A in somatic cells. However, the established link between the activities of Aurora and PLK1, along with the recent work revealing an interplay between Aurora A and PLK4 in forming acentriolar spindle fibres in oocytes strongly suggests that these two kinases interact. The work in this chapter therefore aimed to assess a direct interaction between PLK4 and Aurora A *in vitro,* and an assessment of co-localisation in somatic cells.

6.1.1. **Aims**

The aim of the work described in this chapter was to provide evidence of a direct interaction between PLK4 and Aurora A, and to evaluate the inter-regulation of the catalytic activity of these two enzymes *in vitro* to support the work by Bury *et al*., 2017. Finally, the work in this chapter aimed to confirm an interaction in somatic cells through both co-immunoprecipitation and immunofluorescence experiments and identify at which phase of the cell cycle an interaction occurred.

6.2. Results & Discussion

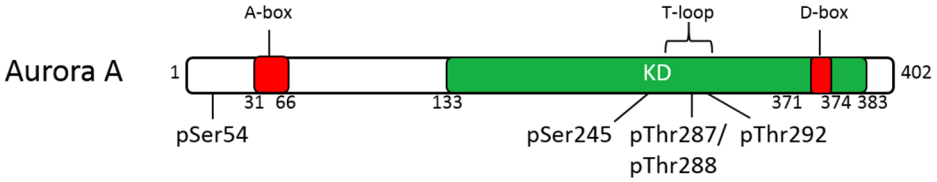
6.2.1. **Aurora A phosphorylation by PLK4**

To determine whether PLK4 phosphorylates Aurora A *in vitro*, recombinant WT PLK4 catalytic domain (1-269) (presented in S. Ferries MRes Report, 2015) was incubated with a recombinantly produced inactive Aurora A, in which the catalytic aspartic acid (D274) was mutated to asparagine to prevent autophosphorylation (made by Dr. Dominic Byrne in the Eyers' laboratory). Phosphorylation of Aurora A observed under these conditions can thus be directly attributed to the activities of PLK4. Following digestion with trypsin, the peptides were analysed by LC-MS/MS, revealing 5 sites of phosphorylation (Figure 6.2 A). One of the phosphosites identified (Ser54) was located at the N-terminus, whilst four were located within the T-loop (Figure 6.2 B). Of particular interest is the identification of phosphorylation at T-loop residue Thr287/288. The sequence surrounding this site conforms to the relaxed PLK4 phosphorylation consensus motif as previously determined using a peptide spot array, which suggests PLK4 preferentially phosphorylates peptides with a basic residue at -3, a Tyr or a hydrophobic residue at +1 and +2 and Ser or Thr at +4 (Johnson *et al.*, 2007; Leung *et al.*, 2007). This therefore suggests a potential role for PLK4 in the regulation of Aurora A activity (Figure 6.2 C).

A

Peptide Sequence	Site of modification	*m/z*	Charge	EThcD/HCD	Mascot Score	Phosphosite confidence
VLCPSNSSQR	S54	614.2	2+	HCD	31	39.42%
VLCPSNSSQRVPLQAQK	S54	665.0	3+	EThcD	35	93.85%
VEFTFPDFVTEGAR	T347	847.9	2+	HCD	21	100%
TATYITELANALSYCHSK	S245	708.3	3+	HCD	54	98.81%
TTLCGTLDYLPPEMIEGR	T292	1073.5	2+	HCD	77	96.95%
(*T*)*T*LCGTLDYLPPEMIEGR	T287/288	1074.0	2+	HCD	60	50%

B



C

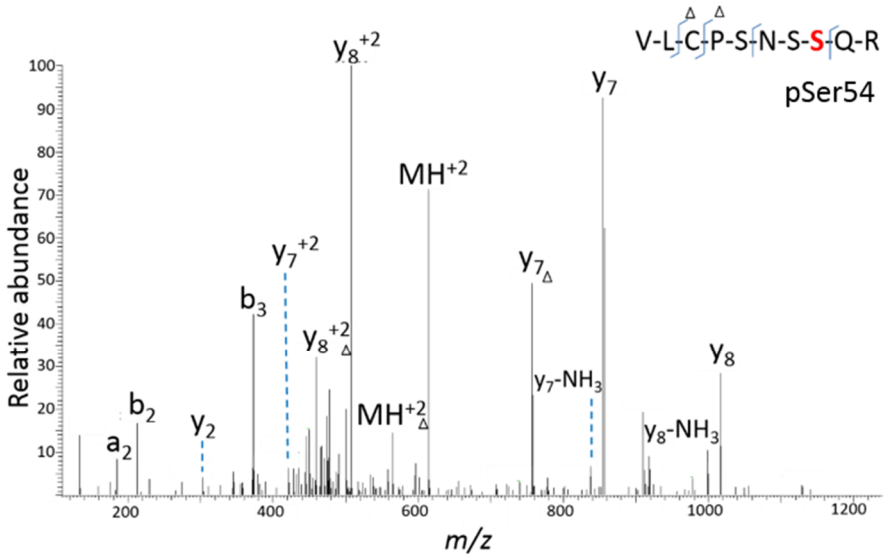
R-X-X-pT-ϕ-X-X-T Relaxed PLK4 consensus

R-R-T-pT-L-C-G-T Aurora A activation loop consensus

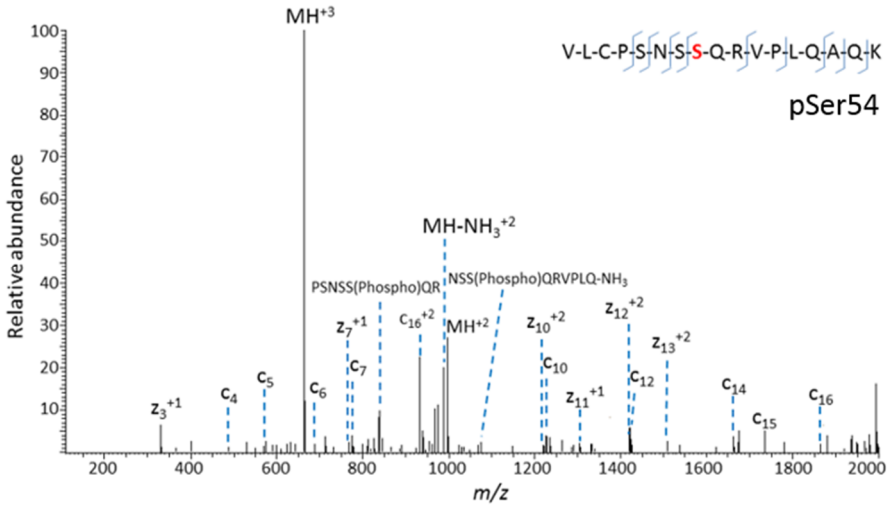
Figure 6.2. WT PLK4 phosphorylates catalytically inactive Aurora A. Purified recombinant WT PLK4 was incubated with catalytically inactive (D274N) Aurora A for 2 hours at 37 °C. Reactions were then digested with trypsin and analysed by LC-MS/MS. A) List of Aurora A phosphorylation sites identified are depicted with associated *m/z* values, Mascot & phosphosite confidence determined using Mascot delta score (%). B) Schematic of Aurora A showing the location of the identified phosphosites. C) Relaxed PLK4 consensus phosphorylation motif and the Aurora A activation loop consensus sequence. ϕ = hydrophobic residue.

Figure 6.3 A to F shows the manual annotation of the product ion mass spectra for each of the identified phosphosites. HCD product ion spectra for the peptide 'VLCPSNSSQR' contains three phosphorylatable residues and insufficient product ions are detected to identify which site is phosphorylated. However, the EThcD spectrum of the same peptide with one missed cleavage provides a *c*- and *z*- ion series, which unequivocally defines Ser54 as the site of modification (Figure 6.3 A & B). Assignment of the HCD product ions in the mass spectrum for the pThr288 containing phosphopeptide identified b_3 and b_5 ions to pinpoint the phosphorylated residue as either Thr287 or Thr288 but the exact site of modification could not be determined (Figure 6.3 D). This ambiguity arises as these two residues are adjacent and located at the extreme N-terminus of the peptide, and the MS/MS spectra lack the y_{17} site determining ion. The inability of collision-induced fragmentation techniques to produce a b_1 ion (due to the lack of a carbonyl group at the N-terminus required for nucleophilic attack to induce peptide cleavage) further prevented the identification of Thr287 or Thr288 as the site of phosphorylation (Schlosser & Lehmann, 2000). Ser245, Thr292 and Thr347 were all identified with high confidence as shown by the manual annotation of HCD product ion spectra (Figure 6.3 C, E & F).

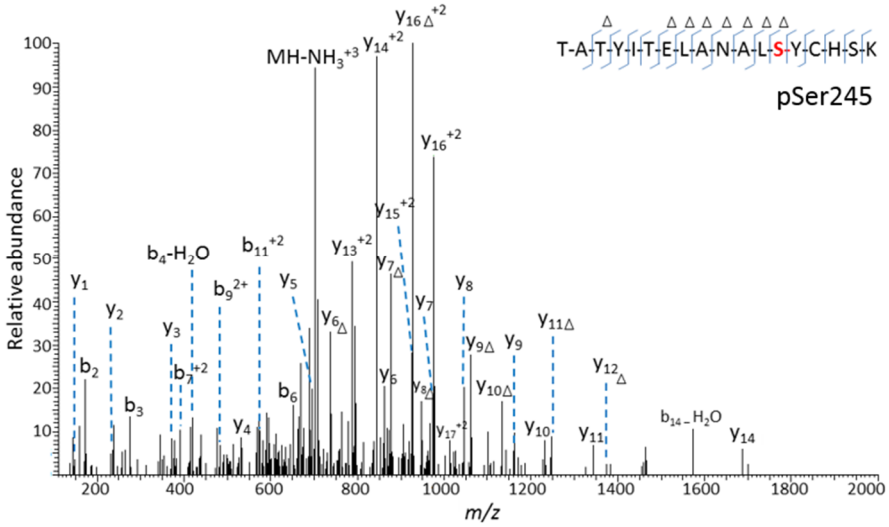
A


V-L-C-P-S-N-S-S-Q-R
pSer54
Relative abundance
m/z
y8+2
y7
MH+2
y7+2
y7Δ
b3
y8+2Δ
y7-NH3
y8
b2
y2
MH+2Δ
y8-NH3
a2

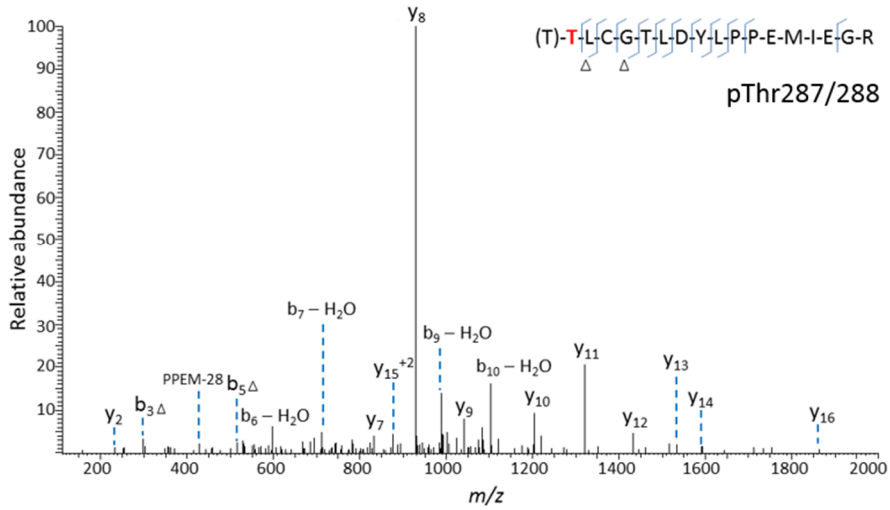
B


V-L-C-P-S-N-S-S-Q-R-V-P-L-Q-A-Q-K
pSer54
Relative abundance
m/z
MH+3
MH-NH3+2
PSNSS(Phospho)QR
NSS(Phospho)QRVPLQ-NH3
z7+1
c16+2
MH+2
z10+2
z12+2
z13+2
c7
c10
c12
z11+1
c14
c16
c5
c6
c4
z3+1
c15

C


T-A-T-Y-I-T-E-L-A-N-A-L-S-Y-C-H-S-K
pSer245
Relative abundance
m/z
MH-NH3+3
y14+2
y16Δ+2
y16+2
y15+2
y13+2
b11+2
b4-H2O
b9 2+
y5
y7Δ
y7
y8
y6Δ
y11Δ
y9Δ
y9
y12Δ
y10Δ
y10
y11
b14 - H2O
y14
b2
b3
b7+2
b6
y4
y8Δ
y17+2
y1
y2
y3

D


(T)-T-L-C-G-T-L-D-Y-L-P-E-M-I-E-G-R
pThr287/288
Relative abundance
m/z
y8
b7 – H2O
b9 – H2O
b10 – H2O
y15+2
y11
y13
y14
y16
y12
y10
y9
y7
b6 – H2O
PPEM-28
b5Δ
b3Δ
y2

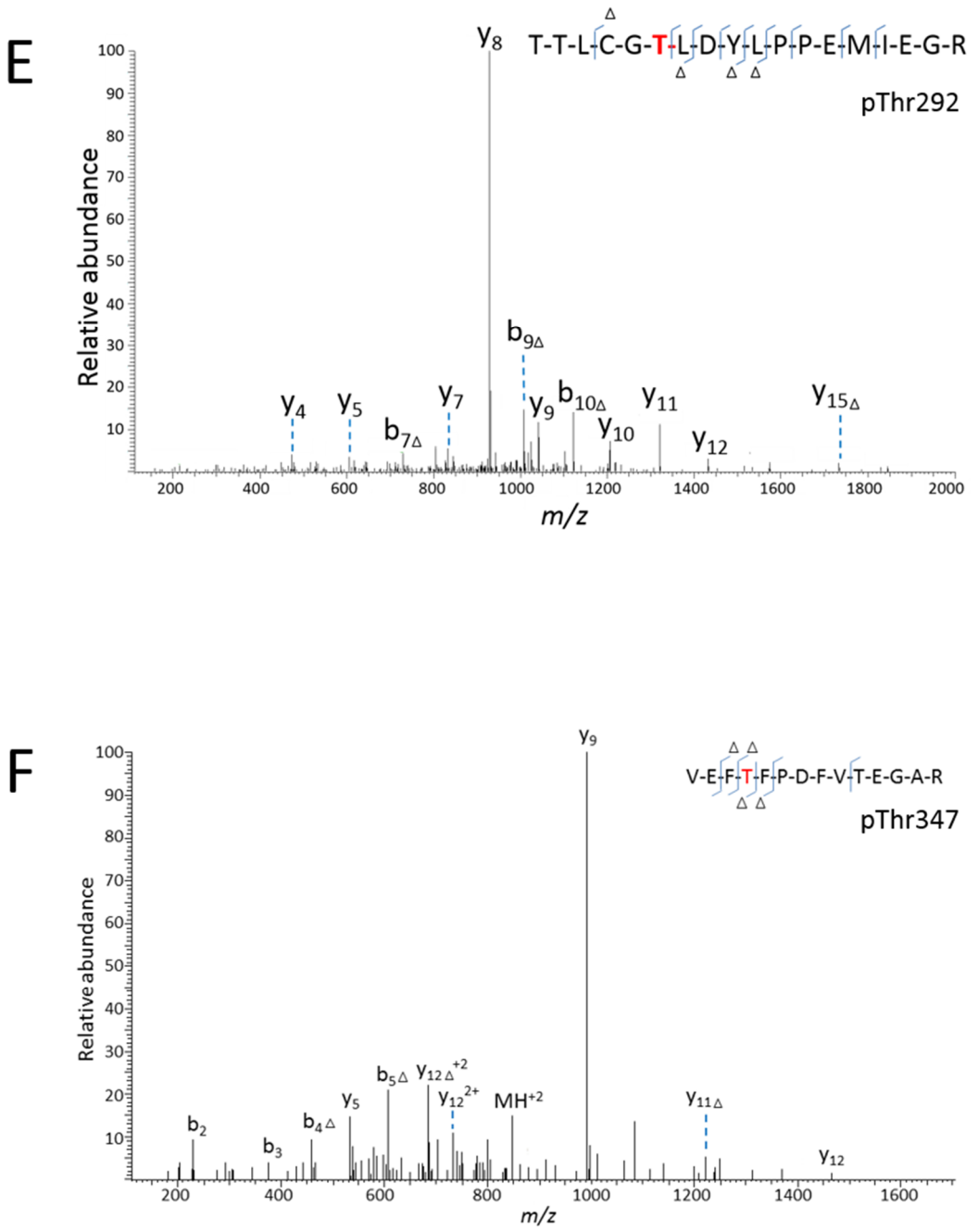


Figure 6.3. PLK4 phosphorylates D274N Aurora A. Tandem mass spectra generated by HCD or EThcD of tryptic phosphopeptides derived from kinase-dead (D274N) Aurora A phosphorylated by PLK4 (1-269). The peptide sequence and the identified phosphosite is detailed on each tandem mass spectrum. A) Doubly charged ion at *m/z* 614.2, generated by HCD suggesting pS54 as a possibly phosphosite; B) Triply charged ion at *m/z* 665.0, generated by EThcD, providing additional evidence for pS54; C) triply charged ion at *m/z* 708.3, generated by HCD, identifying pS245; D) doubly charged ion at *m/z* 1074.0, generated by HCD, identifying the either pT287 or activating T-loop residue, pT288. E) doubly charged ion at *m/z* 1073.5, generated by HCD identifying pT292; F) doubly charged ion at *m/z* 847.9, generated by HCD, identifying pT347.

Owing to the ambiguity surrounding phosphosite localisation at Thr287 or Thr288, further validation was required to confirm that PLK4 phosphorylated Aurora A at the activating residue. Recombinant WT PLK4 (1-269) (presented in S. Ferries MRes Report, 2015) was incubated with catalytically inactive D274N Aurora A over a time course (0 – 2 hours) and proteins were analysed by western blot using a pThr288 specific antibody. This analysis confirmed that PLK4 phosphorylated Aurora A at the critical Thr288 residue in a time-dependent manner. The levels of phosphorylation increase over time, up to 120 minutes and then a decrease in phosphorylation is observed at 240 minutes. This could be due to the proteins becoming unstable and denaturing after incubation at 37 °C for this length of time or possibly due to delayed phosphorylation of Thr287 which would mask that antigenic epitope. Two negative controls were included: a catalytically inactive PLK4 (D154A) incubated with D274N Aurora A, and the inclusion, in one reaction, of the PLK4 specific inhibitor, centrinone. The two negative controls together confirmed that the observed phosphorylation was due to the activities of PLK4 (Figure 6.4).

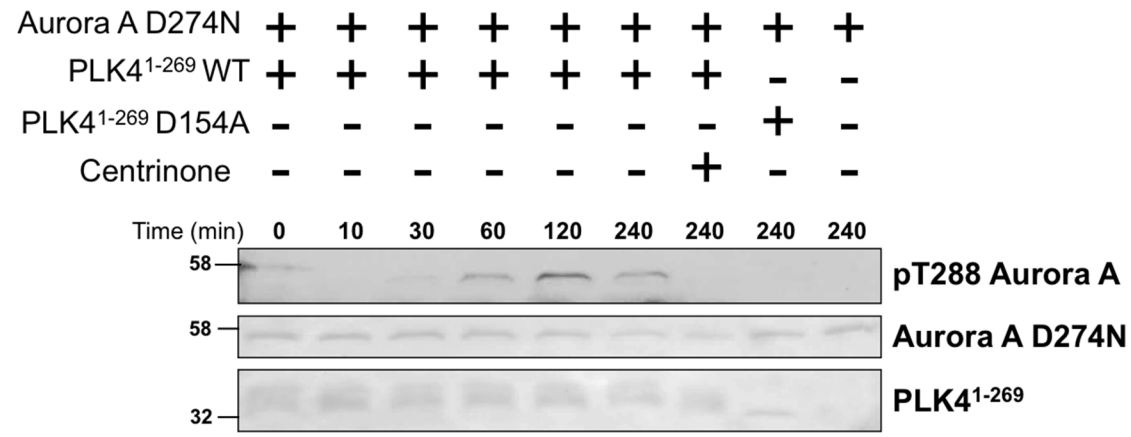


Figure 6.4. PLK4 phosphorylates D274N Aurora A at Thr288. Catalytically inactive Aurora A (D274N) was incubated alone and in the presence of WT PLK4 (1-269), D154A PLK4 (1-269), or in the presence of 10 μM centrinone with 1 mM Mg-ATP. Reactions were analysed at the indicated time points by western blotting using an anti-pThr288 antibody. Equal loading of proteins was confirmed by Ponceau staining.

The identified phosphosites were mapped on the crystal structure of Aurora A using PyMol (Nowakowski *et al.*, 2002). The crystal structure begins at Arg126 and therefore, pSer54 could not be mapped. This site has however, been shown to flank an α-helix motif in the middle of a calmodulin interaction domain and phosphorylation at this site has been implicated in the activities of Aurora A at the basal body in ciliary disassembly and re-absorption during G1/S phase (Plotnikova *et al.*, 2010). Phosphorylation at Ser54 was shown to be important for this process as calmodulin binding to Aurora A resulted in autophosphorylation at Ser54, which is located within the calmodulin binding motif. Mutation of Ser54 prevented binding of Ca^{2+}-calmodulin, preventing mitotic progression and reabsorption of cilia, highlighting the importance of phosphorylation at this residue (Plotnikova *et al.*, 2012).

Phosphorylated Ser245, located on the C-terminal α-helix has been identified as a substrate of GSK3β. This interaction is required to recruit FBXW7, the substrate recognition component of SCF- E3 ubiquitin ligase complex, which mediates the ubiquitination of Aurora, A to signal the protein for degradation (Kwon *et al.*, 2012). Mutation of Ser245 to Ala (along with Ser387) resulted in loss of binding to GSK3β and prevented degradation of Aurora A. This reveals an important regulatory role for phosphorylation at this site, although the implications with regard to possible phosphorylation of this site by PLK4 are unknown.

No functional analysis of phosphorylation at Ser347 has been reported in the literature thus far. Figure 6.5 shows that this residue is located in an unstructured region of the C-lobe. Sequence alignment between different species reveals that this residue is poorly conserved, with lysine more frequently observed at this position (Bayliss *et al.*, 2003). Phosphorylation in intrinsically disordered regions of proteins is common, with the suggestion that phosphorylation in these regions can act as a switch to reveal functionally relevant sites nearby (Iakoucheva *et al.*, 2004; Travers *et al.*, 2015) However, it is possible that the location of this residue in human Aurora A makes it readily accessible for phosphorylation by PLK4 when the two proteins are incubated *in vitro* and may not be biologically relevant.

Figure 6.5 depicts pThr288 and pSer292 on the activation loop, as has been shown previously. No functional analysis of pSer292 has been performed. However, recent work has shown that structurally, when Aurora A is bound to TPX2, Ser292 forms hydrogen bonds with Asp256 of the HRD motif, positioning this aspartic acid to activate the hydroxyl group of the protein substrate. Ser292 therefore appears to aid in the increased activation of TPX2 bound Aurora A (Zorba *et al.*, 2014). The identification of phosphorylated T288 within the T-loop implicates PLK4 activity in the activation of Aurora A and was investigated further.

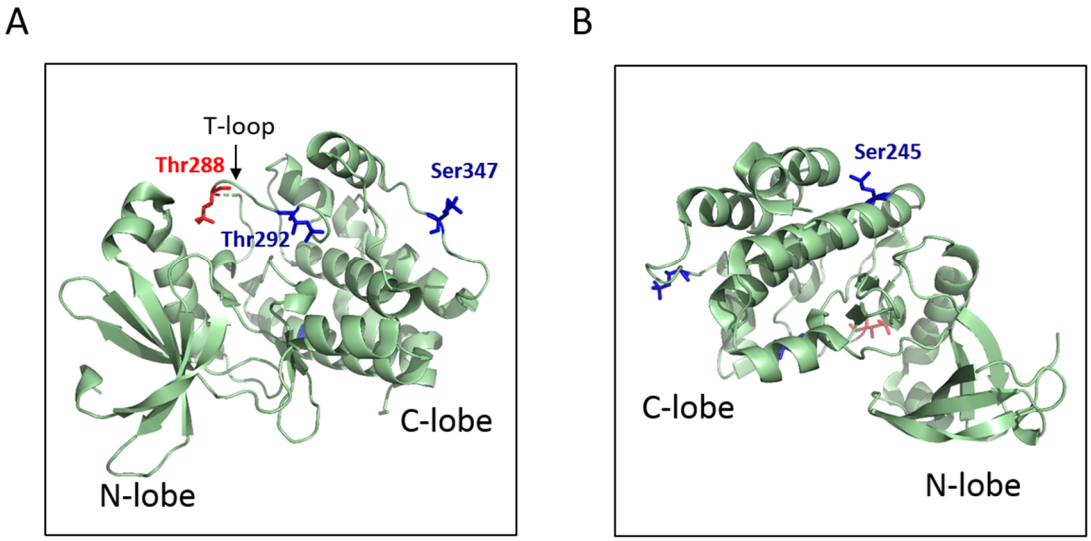


Figure 6.5. Aurora A residues phosphorylated by PLK4. A) The activating residue Thr288 (red) and Thr292 (blue), phosphorylated by PLK4 are located within the T-loop. Ser347 (blue) is located in an unstructured region on the C-lobe. B) Ser245 (blue) is located within a C-terminal alpha helix.

6.2.2. **Aurora A activation by PLK4**

As phosphorylation at Thr288 is critical for Aurora A activity, the effect of phosphorylation at this residue by PLK4 on activity was assessed. Recombinantly produced WT Aurora A is active due to extensive autophosphorylation in *E. coli* and in the presence of ATP *in vitro (Enami & Ishihama, 1984; Shrestha et al., 2012)* (data presented in S.Ferries MRes, 2015). Incubation of Aurora A with a fluorescent peptide substrate (5-FAM-LRRA**S**LG) revealed 30 % conversion of substrate into phosphorylated product after 60 cycles (~75 minutes), with the assay performed at room temperature. However, when WT Aurora A was incubated with WT PLK4 catalytic domain, the activity of Aurora A increased dramatically, leading to a 50 % conversion of the substrate into phosphorylated product (Figure 6.6). The results of this assay confirmed that phosphorylation of Thr288 by PLK4 has a direct impact of Aurora A activity. However, the use of auto-activated Aurora A precluded the early stages of PLK4 activation from being observed. A different approach was therefore required to analyse the onset of Aurora A activation by PLK4.

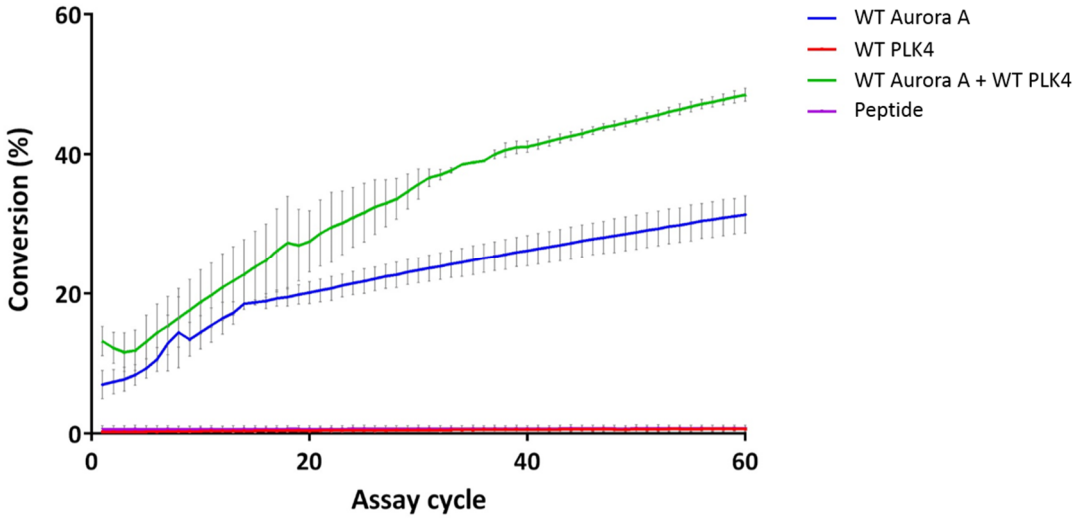


Figure 6.6. Aurora A hyperactivation by PLK4. WT Aurora was assayed for hyperactivation in the presence or absence of WT PLK4 (1-269). Reactions were incubated for 15 minutes at room temperature in the presence of 1 mM ATP to mimic cellular levels prior to evaluating Aurora A activity against the peptide substrate (5-FAM-LRRASLG). Activity was analysed in kinetic mode for 60 cycles (75 minutes) and converted to total peptide phosphorylation. Data are presented as the mean ± SD for triplicate reactions.

To overcome the issue of Aurora A autophosphorylation and activation, 6His-tagged WT Aurora A and 6His- and GST-tagged λ-phosphatase were co-expressed in *E. coli*. Co-expression of these proteins results in the rapid enzymatic dephosphorylation of Aurora A by λ-phosphatase, resulting in a wild type, inactive Aurora A, capable of becoming re-activated upon phosphorylation within its T-loop. The purification steps required to separate the expressed Aurora A from λ-phosphatase are shown (Figure 6.7, taken from S. Ferries MRes Report, 2015). Following nickel affinity and glutathione S-transferase purification, the flow through (expected to contain purified Aurora A) contains a contaminant corresponding to the molecular weight of 6His- GST tagged λ-phosphatase (Figure 6.7 A, B). Therefore, strong cation exchange chromatography was employed to exploit the differences in PI between the two proteins (pI of Aurora A = 9.5; pI of λ-phosphatase= 4.5) as a final purification step. Fractions 1-7 (Figure 6.7 C) were pooled and analysed by LC-MS/MS to confirm the identity of dephosphorylated WT Aurora A (data not shown).

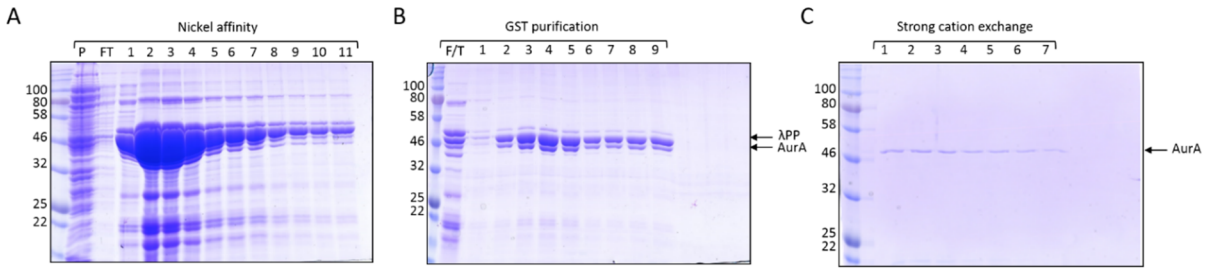


Figure 6.7. Purification of dephosphorylated WT Aurora A. A) Following co-expression of 6His-tagged Aurora A & 6His-GST λ-phosphatase in Rosetta *E. coli*, cell lysate was incubated with nickel resin and bound protein was eluted with 500 mM imidazole. Eleven fractions were collected and analysed by SDS-PAGE. B) Fractions 1-11 were pooled and Aurora A was separated from λ-phosphatase by GST affinity purification. Nine fractions were collected and analysed by SDS-PAGE. C) The flow through from B) was purified further by strong cation exchange and collected fractions were analysed by SDS-PAGE. Fractions 1-6 were pooled and analysed by LC-MS/MS to confirm protein identity.

To assess the early stages of Aurora A activation by PLK4, the purified, dephosphorylated Aurora A was incubated with and without WT PLK4 (1-269) for 2 hours at 37 °C. The kinase-inactive D154A PLK4 (1-269) mutant was incubated with Aurora A as a negative control and GST-tagged TPX2, the allosteric activator of Aurora A, was included as a positive control. Figure 6.8 shows the onset of activation of dephosphorylated Aurora A, with low activity toward the peptide substrate. A maximum of 3.5 % conversion of the substrate was observed after 60 assay cycles. The overall activity was much lower than that observed in Figure 6.6 and this may be due to a slow rate of reaction or due to the dephosphorylated Aurora A having an altered conformation. Phosphorylation regulates protein stability in many instances and therefore, a dephosphorylated protein kinase may be less stable and therefore less active (Nishi *et al.*, 2014). However, a clear increase in the activity of Aurora A against the peptide substrate is observed when incubated with PLK4 and is similar to that observed with TPX2 (10 % vs 12 % respectively), representing a ~3 fold increase in activity. To confirm that the effect observed with TPX2 was not due to the presence of a GST tag, GST alone was incubated with Aurora A as an additional negative control.

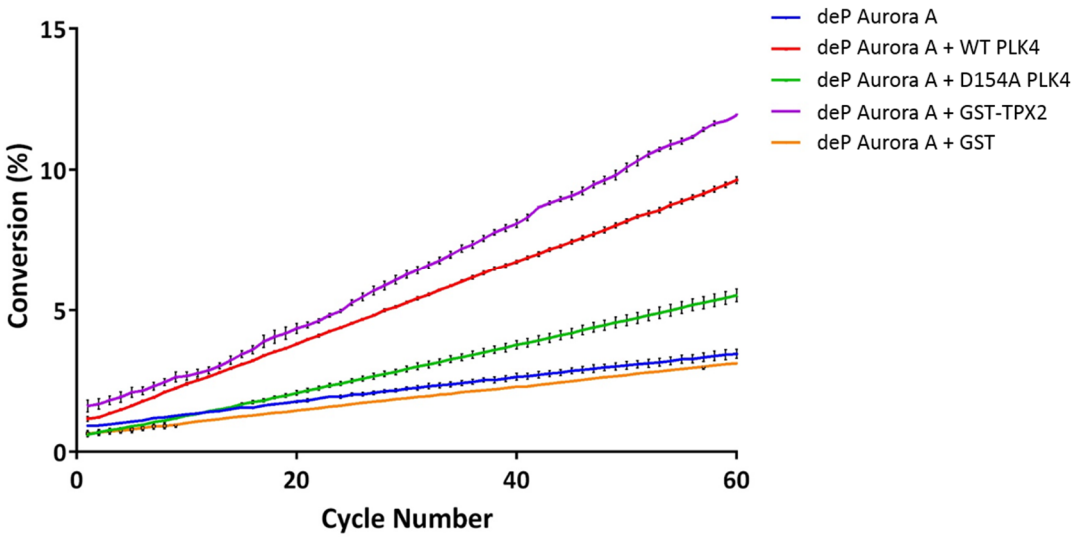


Figure 6.8. Dephosphorylated Aurora A activation by WT PLK4. λ-phosphatase treated WT Aurora was assayed for WT PLK4, catalytically inactive D154A PLK4, or TPX2-mediated hyperactivation. Reactions were incubated for 30 minutes at 37 °C in the presence of 1 mM ATP to mimic cellular levels. Aurora A activity was assayed upon the addition of the peptide substrate (5-FAM-LRRASLG) in kinetic mode for 60 cycles (75 minutes) and converted to total peptide phosphorylation. Data are presented as the mean ± SD for triplicate reactions.

6.2.3. **Aurora A & PLK4 co-immunoprecipitation**

Taken together, the data show that *in vitro*, PLK4 phosphorylates Aurora A and that phosphorylation within the T-loop leads to an increase in activity. Whilst these data are encouraging, validation in cells is required to confirm that this interaction is biologically relevant. To test this, co-immunoprecipitation experiments were performed to assess whether PLK4 and Aurora A co-purified from cells. Expression of a FLAG-tagged Aurora A and transiently transfected MYC-PLK4 was induced and Aurora A was immunoprecipitated by virtue of its FLAG tag. Analysis by western blot shows the successful transient transfection and expression of MYC-PLK4 alongside the expression of FLAG-Aurora A in the total cell lysate. However, following immunoprecipitation of Aurora A, no PLK4 was observed in the pull-down, suggesting that Aurora A and PLK4 do not co-immunoprecipitate, at least under the conditions employed (Figure 6.9 A).

As an alternative approach, co-immunoprecipitation was tested in FLAG-PLK4 stably transfected U2OS cells. FLAG PLK4 was immunoprecipitated by virtue of its FLAG tag and co-purification of endogenous Aurora A was assessed. Owing to the low expression levels of stably transfected PLK4, the proteasome inhibitor MG-132 was included to permit the accumulation of PLK4 levels in the cell. Figure 6.9 B shows the enrichment of PLK4 following immunoprecipitation. However, co-purification with Aurora A was not observed.

Taken together, Figure 6.9 A & B show no evidence of interaction between PLK4 and Aurora A in cells therefore, an alternative approach was required to investigate a possible interaction.

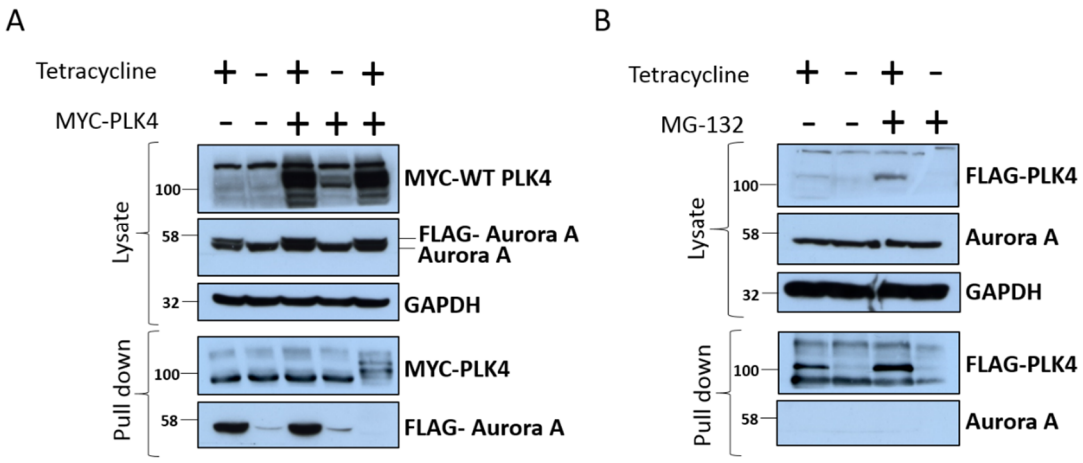


Figure 6.9. PLK4 & Aurora A do not co-immunoprecipitate. A) HeLa cells stably transfected with WT Aurora A were co-transfected with MYC- WT PLK4 and incubated +/- 1 µg/mL tetracycline for 18 hours to induce protein expression. Aurora A was immunoprecipitated with anti-FLAG resin. Lysates and pull down samples were analysed by western blot and probed using the indicated antibodies. B) U2OS cells stably transfected with FLAG-PLK4 were incubated +/- 1 µg/mL tetracycline for 18 hours to induce protein expression followed by +/- 10 µM MG-132 for 4 hour. PLK4 was immunoprecipitated with anti-FLAG resin. Lysates and pull down samples were analysed by western blot and probed using the indicated antibodies

6.2.4. **Aurora A & PLK4 co-localisation**

As an alternative to co-immunoprecipitation, U2OS cells expressing MYC-PLK4 were imaged by immunofluorescence to assess whether PLK4 and Aurora A co-localised in the cell. Figure 6.10 shows WT PLK4 co-localised with Aurora A within the nucleus, in an interphase cell, as indicated by the decondensed DNA. The presence of two puncta reflects that the cell has duplicated its centrosome and so is either in S phase or progressed in to G2. The exact phase cannot be determined without imaging markers distinct to different phases of the cell cycle. This assay would therefore need to be repeated and fluorescently probed with a cell cycle marker such as Cyclin B1 to determine whether this interaction has occurred following progression in to G2 (Scott *et al.*, 2003).

Aurora A could be observed in cells across the cell cycle. However, as expected, co-localisation with PLK4 was not observed in cells in the later stages of mitosis. PLK4 trans-autophosphorylates and is degraded by the 26S proteasome following centriole duplication and is not believed to function in mitosis. Studies investigating the subcellular localisation of Aurora A have shown that low levels of Aurora A can be observed in G1 which then accumulate at duplicated centrosomes from late S phase through mitosis, which is consistent with the data shown here (Dutertre *et al.*, 2002).

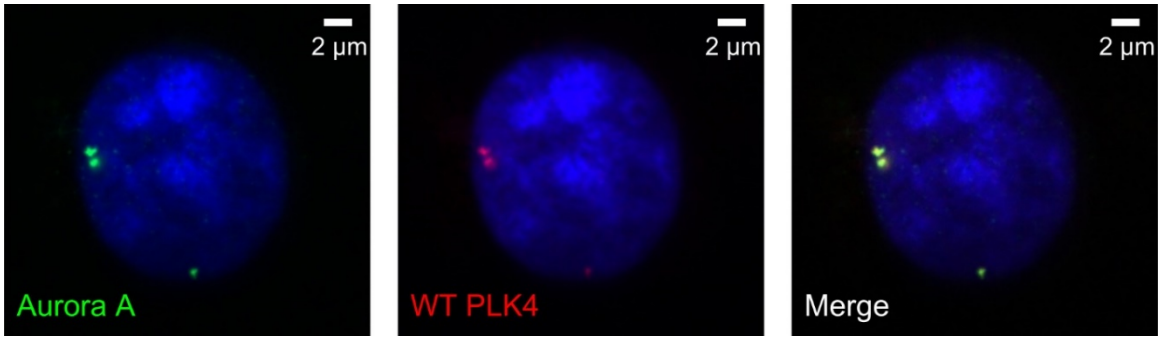


Figure 6.10. PLK4 and Aurora A co-localise in U2OS cells. U2OS cells, induced with tetracycline to express MYC-WT PLK4, were probed with anti-PLK4 (red) & anti-Aurora A (green) antibodies and analysed by immunofluorescence microscopy to confirm co-localisation.

Next, co-localisation of PLK4 with active, Thr288 phosphorylated Aurora A was assessed. Using a pThr288 specific antibody, MYC-WT PLK4 expressing U2OS cells were imaged by immunofluorescence. Cells at varying stages of the cell cycle were imaged to identify at which stages phosphorylated Thr288 could be identified and possibly interact with PLK4. Figure 6.11 shows that pThr288 Aurora A can be identified at each of the stages of the cell cycle indicated, consistent with the literature (Walter *et al.*, 2000). Figure 6.11.B shows pThr288 Aurora A within a metaphase cell. Active Aurora A (pThr288 phosphorylated) has recently been shown to play an unexpected role in kinetochore attachment in metaphase cells when the chromosomes reside near spindle poles. This occurs via phosphorylation of the Hec1 subunit from the NDC80 complex, which associates with the kinetochore to control attachment to the microtubules (DeLuca *et al.*, 2018). Figure 6.11 C shows a cell in late anaphase/telophase, where active Aurora A can be identified at the spindle mid-zone and centrosomes, which has been shown previously (Marumoto *et al.*, 2003). This reflects the roles of Aurora A in the later stages of the cell cycle, including the completion of cytokinesis. As expected, co-localisation with PLK4 was not detected in the cells at later stages of the cell cycle due to PLK4 expression being tightly regulated in interphase cells.

However, co-localisation of pThr288 Aurora A with PLK4 was observed in the early stages of the cell cycle (G1/S). Figure 6.11 A shows two cells that have recently undergone cytokinesis and re-entered G1. In these two cells, low levels of pThr288 Aurora A can be observed as single puncta, co-localised with PLK4. Previous studies showing Thr288 phosphorylation across the cell cycle revealed that pThr288 levels increase from G1 through to G2 and then decrease during mitosis, before elevating again upon re-entry in to

G1 (Walter *et al.*, 2000). This suggests an active role for pThr288 Aurora A in the earlier phases of the cell cycle (G1/S) and the data shown here suggests that this activity could potentially be regulated by the catalytic activity of PLK4.

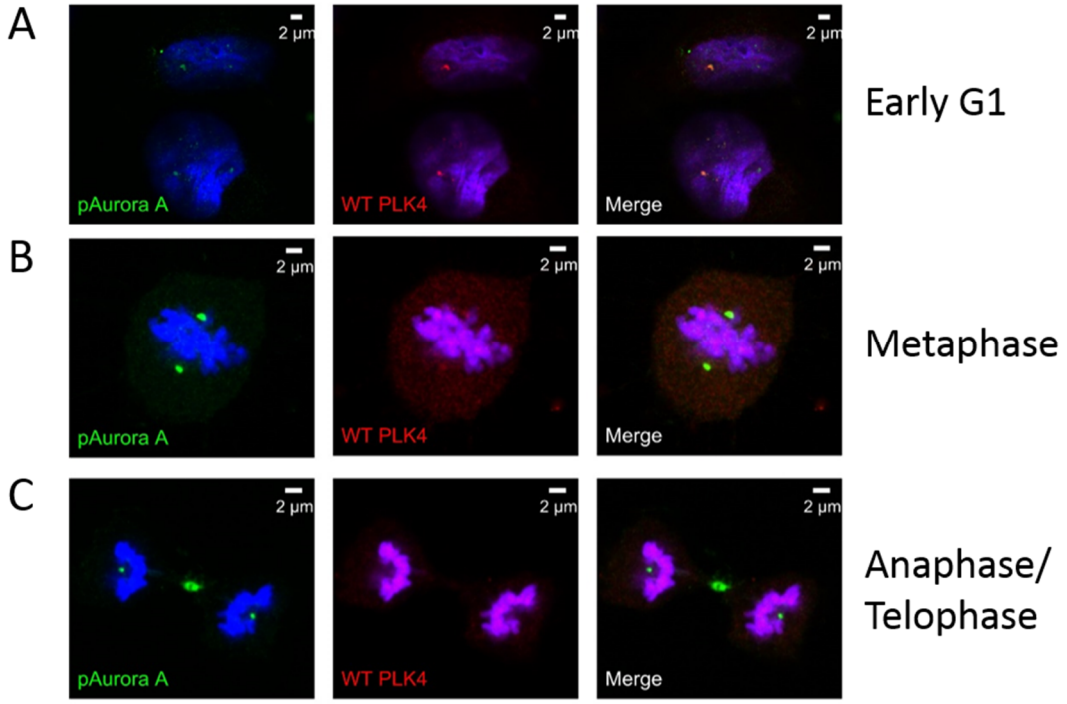


Figure 6.11. PLK4 and pT288 Aurora A co-localise in U2OS cells. U2OS cells, induced with tetracycline to express MYC-WT PLK4, were probed with anti-PLK4 (red) & anti-pThr288 Aurora A (green) antibodies and analysed by immunofluorescence microscopy to assess co-localisation in cells A) early G1 phase B) metaphase C) late anaphase/telophase.

6.3. Conclusions

This chapter describes the attempt to identify an interaction between PLK4 & Aurora A for the first time, demonstrating that PLK4 acts upstream of Aurora A, regulating its activity by phosphorylation. This work showed that PLK4 is capable of phosphorylating Aurora A *in vitro* at 5 sites, including Thr288, the activating residue located within the T-loop. Of the 5 observed phosphosites, only Thr288 has been studied extensively to understand its biological function, with two other phosphosites (Ser52 and Thr292) having been observed in high-throughput proteomic analysis, but lacking information on functional roles. Future studies should therefore aim to understand the possible biological relevance of these PLK4 directed phosphosites through approaches such as chemical genetics to mutate these residues and transfect in to cells and observe the phenotypic effect. This approach has been used frequently in the study of protein kinases, including PLK4 and Aurora A to probe their biological functions (Scutt *et al.*, 2009; Holland *et al.*, 2010; Sloane *et al.*, 2010).

Following confirmation that PLK4 phosphorylates Aurora A at Thr288, *in vitro* kinase assays showed that this phosphorylation resulted in hyperactivation of Aurora A. Given the known complexity of Aurora A activation through regulation by multiple cofactors (TPX2, Bora, Ajuba), it is possible that this work has revealed a novel regulatory mechanism of Aurora A activation, via phosphorylation by PLK4.

The immunofluorescence images indicated that the co-localistion of the two kinases occurred only at earlier stages of the cell cycle (G1/S) where potential roles of pThr288 Aurora A have yet to be elucidated. This could provide an exciting opportunity to uncover novel roles for Aurora A in the earlier stages of centriole biogenesis.

Although PLK4 and Aurora A co-localised in the cell, co-purification could not be achieved. Co-immunoprecipitation experiments are challenging and require a level of optimisation to ensure conditions are suitable to co-purify two interacting proteins. Factors such as buffer components (pH, salt concentration, detergent concentration) and total number of washes performed can affect the stability of an interaction (Masters, 2004). In addition, it is unlikely that all of the protein expressed is bound to just one interacting partner and the interactions that do occur are likely to be transient and occur only at distinct phases of the cell cycle. This will further decrease the probability of observing an interaction by techniques such as western blotting.

Overall, this work has identified Aurora A as a possible substrate of PLK4 *in vitro* and revealed co-localisation at early phases of the cell cycle, supporting the findings of Bury *et al*., which showed overlapping functions of PLK4 and Aurora A in forming acentriolar spindles in oocytes. This work, along with the *in vitro* and U2OS cell-based experiments described in this chapter, poses the possibility that in addition to a role in meiosis in oocytes, PLK4 and Aurora A may also interact in somatic cells to co-ordinate activities in the early stages of the cell cycle. This preliminary work provides evidence of a direct interaction. However, a more comprehensive study of this interaction in cells is required to unravel the possible interplay of these two important mitotic protein kinases.

7. General Discussion and Future Perspectives

The work in this thesis aimed to uncover PLK4 signalling and identify novel PLK4 substrates. PLK4 is the master regulator of centriole biogenesis, with tightly controlled activity to ensure that each centriole is duplicated once per cell cycle. Despite its critical importance, few PLK4 substrates have been identified thus far, and much remains to be understood about its role at the centrosomes and its wider role regulating activities in the cell. The work in this thesis has revealed a number of novel PLK4 regulated signalling pathways, identified a new PLK4 phosphorylation consensus sequence and confirmed PLK4 as an activator of Aurora A activity.

The first results chapter (Chapter 3) described the development of biochemical tools to study the PLK4 regulated phosphoproteome, including characterisation of recombinant WT and G95R PLK4, and assessment of their inhibition by centrinone. This work confirmed the drug-resistant phenotype *in vitro* prior to analysis of both proteins in cells. Chapter 3 also described the successful creation of stable isogenic U2OS cell lines, transfected with either WT or G95R PLK4, where protein expression and correct subcellular localisation was confirmed. This was critically important, as the work in this thesis aimed to study the activities of PLK4 in the cell and therefore, the overexpression system had to mimic endogenous conditions as closely as possible. Finally, the work in Chapter 3 assessed the efficiency of SILAC labelling in U2OS cells to confirm the cell doublings required for full incorporation of the labels, and established the high pH reversed phase fractionation strategy to be employed for improved (phospho)peptide coverage. Together, this work permitted a comprehensive analysis of the PLK4 regulated phosphoproteome.

Chapter 4 details how mass spectrometry acquisition methods were optimised for large-scale confident phosphopeptide and phosphosite analysis. A number of instrument and informatics parameters were tested, with a key focus on mode of fragmentation and mass analysers for product ion detection. HCD fragmentation, with detection of both precursor and product ions in the orbitrap was identified as the optimal method for identifying high numbers of phosphopeptides, with high confidence in position of the phosphosite. Whilst EThcD fragmentation with product ion detection in the orbitrap provided the highest phosphosite confidence compared with HCD (orbitrap or ion trap detection), the total number of phosphopeptide identifications was lowest. This is likely due to the longer

reaction times required for efficient ETD fragmentation. Further advances in ETD technology are required for it to achieve the same numbers of identified phosphopeptides compared with HCD in large-scale phosphoproteomics studies, including improved reagent ion sources and advances in supplemental activation. Currently there is a balance between numbers of phosphopeptides identified and numbers of phosphosites confidently localised which needs to be taken into consideration when designing an analytical strategy for phosphoproteomics.

Recent improvements with ETD supplemental activation have included the development of an activated ion- negative ETD (AI-NETD) fragmentation approach for peptides. Negative mode ETD, in which peptides are oxidized with reagent cations leading to cleavage of the C-Cα bond has been shown to produce interpretable MS2 spectra for peptide identification. However, product ions can form intramolecular non-covalent interactions that do not separate (McAlister *et al.*, 2012; Riley *et al.*, 2015). In AI-NETD, infrared photon bombardment occurs at the same time as the ETD reaction, resulting in dissociation of the non-covalent interactions and richer MS2 spectra. Combined with a multipurpose dissociation cell implemented in a quadrupole ion trap-orbitrap hybrid mass spectrometer, this negative ETD fragmentation method identified ~7600 peptides in a single run from a tryptic digest of *Saccharomyces cerevisiae* (Riley *et al.*, 2015). As phosphopeptides preferentially ionize to form anions, future developments with this approach could lead to the ability to implement AI-NETD for large-scale phosphoproteomic studies.

The implementation of ETD for high-throughput analysis will also benefit from practical considerations such as improved reaction cells for optimised ETD reactions. Improvements in this area could allow more efficient storage of precursors and reagent ions and reduced ion-ion reaction times. In addition, concurrent supplemental activation as opposed to two separate reactions would lead to faster reaction times and permit the acquisition of more ETD spectra leading to larger numbers of phosphopeptide identifications (Riley & Coon, 2018).

Following on from the development of biochemical tools and optimised MS parameters, a large-scale phosphoproteomics investigation was performed to study PLK4-regulated signalling via the analysis of differential protein expression and changes in phosphorylation state upon inhibition of PLK4 activity. This work revealed that PLK4 was implicated in regulating MAPK signalling and the early phases of the cell cycle. In addition, PLK4 activity

was implicated in roles outside of the cell cycle, including rRNA processing, cadherin binding involved in cell-cell adhesion, and negative regulation of apoptosis.

A number of dysregulated phosphoproteins were also identified as enriched at the centrosome. As the centrosome cycle is strictly linked to DNA replication at S phase, it is not surprising that downregulated proteins/phosphosites were observed with roles in both of these processes. CDKs are known to play a role in mediating centrosome biogenesis and the downregulation of CDK4 and 6 with roles in both G1/S and at the centrosome highlights this link. A gene knockout approach confirmed the importance of CDK4 in the centrosome cycle as $Cdk4^{-/-}$ MEFs failed to undergo centrosome duplication (Adon *et al.*, 2010). Similarly, RB (identified with up and down regulated phosphosites) is also involved in the centrosome cycle, as release of E2F following RB phosphorylation leads to the regulation of genes involved in centriole duplication, including cyclin D, cyclin A, CDK2 and RanBP (Ishida *et al.*, 2001; Ren *et al.*, 2002). This work may have therefore provided further evidence of the cross-talk between proteins regulating both centrosome duplication and progression through G1/S phase of the cell and showed them to be regulated by PLK4 signalling.

As PLK4 is localised at the centrosome, a more focused approach to studying PLK4 signalling could be to isolate and analyse the centrosomal phosphoproteome. Subcellular phosphoproteomics permits a greater insight in to the intricate regulation of phosphorylation dynamics and isolation of one subcellular compartment allows for lower abundance proteins to be identified (Trost *et al.*, 2010). This approach was taken to study the PLK1 regulated phosphoproteome of the mitotic spindles. Compared to analysis of the global proteome, this study provided a comprehensive assessment of regulated proteins at the spindle fibres, and identified a number of novel PLK1 regulated phosphosites on spindle associated proteins such as INCENP and Bub1 (Santamaria *et al.*, 2011).

In addition, future work could include arresting the cells in G1 phase of the cell cycle and analysing the regulation of proteins/phosphosites with roles in the early phases of the cell cycle where PLK4 is most active. Proteomics datasets can also be coupled to RNA-sequencing data to assess the effects of drug-treatment on transcript levels. This was performed for a large-scale study, which aimed to identify mediator kinase substrates using the specific CDK8 and CDK19 inhibitor cortistatin A. RNA-sequencing data correlated with the targets identified in the proteomics study and provides additional depth in the

understanding of kinase inhibition at the transcript, protein and PTM level (Poss *et al.*, 2016).

These approaches to studying the activity of PLK4 at the centrosome could aid the identification of increased numbers of PLK4-regulated centrosomal proteins, including the known PLK4-regulated phosphosites in STIL, CP110 and Cep192/152 not seen in this work.

One noteworthy observation is that the fold changes observed in the phosphoproteomics study were typically relatively small (<2 fold). Many large-scale phosphoproteomics studies implement an arbitrary fold change cut off (typically 2-fold) to determine significance and employ a t-test to assess statistical significance of the calculated ratios for SILAC data. The t-test has a number of limitations when used in proteomics studies. Firstly, the sample size in this type of experiment is typically small, (3 to 5 bioreplicates) and this results in issues in estimating the true variability in the data. This is exacerbated for phosphoproteomics datasets, where SILAC ratios determined for single peptides, compared with protein quantification where a minimum number of peptides are required for quantification. As the variability estimates are used in the t-test statistic, proteins and peptides can be declared as not significant due to high variability.

Empirical Bayesian approaches have been the gold standard for gene microarray data, although its use in the proteomics field has lagged behind. Bayesian statistics allow for a realistic distribution of variance by using the full dataset to shrink the variances observed toward a pooled estimate (Kammers *et al.*, 2015). In this way, Bayesian statistics can determine statistical significance, even with small sample sizes and was therefore an ideal solution to determine significance in the phosphoproteomics dataset described in Chapter 5, where variability in phosphopeptide quantification and small fold changes were observed.

It is possible that the small fold changes observed were due to the short centrinone incubation times (4 hours). This was likely to be insufficient to study the PLK4-regulated proteome due to differential protein turnover rates (Doherty *et al.*, 2009). Similarly, a longer incubation time may be required to increase the effects of inhibited PLK4 activity at the level of phosphorylation. Therefore, future studies could include assessment of the PLK4 regulated phosphoproteome following centrinone treatment for longer time periods.

An additional complication of phosphoproteomics studies arises from the consideration of the protein and phosphopeptide datasets as separate. Ideally, the phosphosite SILAC ratios would be normalised to the total protein ratios, to permit the determination of stoichiometric changes in phosphorylation. However, to date, there is no established method for this. One major complication is that there will be many instances where a phosphopeptide has been identified, with no corresponding protein identification, which would lead to a loss of potentially important data. This work therefore considered the two datasets separately; referring to the protein level data in instances where the protein and a phosphosite were both identified as significantly changing following treatment with centrinone. Future advances in instrumentation will undoubtedly lead to even greater numbers of peptide identifications, maximising coverage of the proteome under study and potentially overcoming the issue of missing data currently preventing normalisation.

One exciting result described in this thesis was the identification of a potential proline directed PLK4 phosphorylation consensus motif. This motif was validated using *in vitro* kinase assays, which showed PLK4 phosphorylation of both SP and PSP directed motifs. These data were somewhat unexpected, as this consensus motif varies significantly from that proposed from peptide spot arrays, in which proline at +1 was not tolerated (Leung *et al.*, 2007). In addition, proline-directed motifs are known to be targets of cyclin-dependent kinases (Songyang *et al.*, 1994; Swaffer *et al.*, 2016), and a number of the downregulated phosphosites identified are direct substrates of CDK activity.

NCKAP5L, discussed in Chapter 5, has also been identified as a substrate of CDK1. Phosphorylation within a proline rich domain of NCKAP5L by CDK1 at seven residues results in dissociation from the centrosome during mitosis (Mori *et al.*, 2015b). The sites included Ser571 and Ser577 however, Ser567, identified in this study, was not observed. This suggests that both PLK4 and CDKs can phosphorylate NCKAP5L within the proline directed motifs, and additional residues nearby may dictate subtle differences in their target proteins. For example, Lys/Arg at the +3 position is included in the full CDK consensus motif (Songyang *et al.*, 1994).

Cdc6 pSer74, identified as downregulated in this phosphoproteomics study in response to centrinone treatment, is located near to a KEN destruction box and is a known target of CDKs. CDK2-cyclin E phosphorylation of Ser74 stabilises Cdc6 by preventing its degradation by the APC/C E3 ubiquitin ligase and allows pre-replicative complexes to assemble prior to S

phase (Mailand & Diffley, 2005). Following the completion of G1/S phase, CDK phosphorylation of Ser74 causes Cdc6 to be exported out of the nucleus in to the cytoplasm to prevent the formation of pre-replication complexes (Yim *et al.*, 2013). This suggests that downregulated Cdc6 pSer74 in this study is an off-target effect of centrinone treatment.

A number of downregulated phosphosites have also been shown to be substrates of Aurora A. WDR62 phosphorylation at Ser33 is a target of TPX2- bound Aurora A to recruit WDR62 to the mitotic spindles with roles in organisation and chromosome alignment during metaphase (Lim *et al.*, 2015). Similarly, Aurora A phosphorylates NUMA1 at Ser1991, which has roles in spindle orientation (Gallini *et al.*, 2016).

The identification of downregulated phosphorylation of established CDK & Aurora A targets suggests the possibility of off-target inhibition of these proteins with centrinone. As shown in Chapter 3, centrinone is capable of inhibiting Aurora A, and so it is possible that centrinone can also inhibitor CDKs, as dual inhibition of Aurora A, and CDKs has been shown with the compound JNJ-7706621 (Emanuel *et al.*, 2005). Taken together, these observations highlight the complexity of using a small molecule inhibitor to study the activities of a specific protein kinase and the requirement for methods to validate the results. Downregulated phosphosites within PSP motifs of target proteins will need to be validated as true PLK4 substrates by other methods. This could include *in vitro* kinase assays, in which PLK4 is incubated with potential substrates identified in the phosphoproteomics study. Subsequent LC-MS/MS analysis could then be used to identify phosphorylated peptides containing the PSP motif.

This work attempted to validate PLK4 targets by the parallel analysis of a drug-resistant G95R PLK4 expressing cell line, to monitor the off-target effects of centrinone. Analysis of the G95R PLK4 expressing cells revealed a significant number of dysregulated proteins and phosphosites, with the strong possibility that the use of a drug-resistant mutant led to an increased amount of unbound centrinone in the cell, which could then bind to off-target proteins with higher affinity. Unfortunately, this meant that the dataset could not be used to validate true targets of WT PLK4 signalling. Incubating the WT PLK4 expressing cells with higher concentrations of centrinone and analysing the effects on the phosphoproteome would allow for the off-target effects of centrinone to be studied. If the dysregulated proteins/phosphosites in the G95R PLK4 dataset are due to off-target effects, then the same should be observed in the WT PLK4 expressing cells treated with a higher

concentration of centrinone. An additional limitation of the G95R PLK4 dataset was the presence of endogenous PLK4. Whilst expressed at very low levels in the cell, it is still possible that a number of the regulated proteins/phosphosites in the data were due to inhibition of the endogenous protein.

There is a pressing need to find solutions to validate phosphoproteomics datasets acquired following inhibition of a protein kinase with a small-molecule inhibitor. One solution could be to abolish kinase activity by protein depletion using an siRNA approach. This would increase the specificity of the response and allow dysregulated proteins and phosphosites to be linked directly to PLK4 activity with greater confidence (Santamaria *et al.*, 2011). Additionally, target proteins need to be validated by western blotting where specific antibodies are available. Chapter 5 showed downregulation of cyclin D2 following centrinone treatment of WT PLK4 expressing cells to match the proteomics datasets. However, the decrease observed was small and therefore, a centrinone time-course could be employed to assess down regulation of proteins over time. This was performed to assess EGFR and pThr669 EGFR upregulation however; conflicting results were obtained, with no consistent changes in expression observed. This data reflects the complexity of cell signalling mechanisms within the cell and further highlights the requirement for robust validation of phosphoproteomics data.

Chapter 6 described an interaction between PLK4 and Aurora A, whereby PLK4 phosphorylated Aurora A, including within the activation loop, leading to hyperactivation. Analysis of the interaction in cells confirmed that both Aurora A and pThr288 Aurora A localise with PLK4 in interphase cells. However, this work was unable to show that the two proteins co-immunoprecipitated. Using the knowledge that PLK4 and Aurora A appear to co-localise in interphase cells, an alternative approach to study this interaction could include arresting the cells in G1/S and repeating the co-immunoprecipitation experiment. In this way, all of the cells should remain in interphase where PLK4 is most abundant and active, and increase the likelihood of observing an interaction. In addition, the role of PLK4 phosphorylated pThr288 Aurora A in the early phases of the cell cycle needs to be understood to ascertain the functional relevance of this interaction in somatic cells. This is a complicated task, as pThr288 is also a target of autophosphorylation and can be enhanced upon binding to co-factors such as TPX2 (Eyers *et al.*, 2003). Therefore, unravelling the separate roles of pThr288 resulting from autophosphorylation or PLK4 will be very challenging. As the interaction was identified in a G1/S phase cell, it is possible that PLK4

regulated phosphorylation of Aurora A has a functional role in centrosome duplication, whereas TPX2 bound, pThr288 autophosphorylated Aurora A has roles in the later stages of the cell cycle at the mitotic spindle.

Overall, the work in this thesis described the development of biochemical tools to study PLK4 in cells, and the optimisation of MS acquisition parameters and data processing, which allowed for a comprehensive investigation in the PLK4-regulated phosphoproteome for the first time. This work revealed novel PLK4-regulated proteins and phosphosites implicated in a variety of cellular processes, including the cell cycle, rRNA processing and regulation of apoptosis. In particular, the identification of a potential novel PLK4 consensus sequence will prove a useful starting point for future work aiming to investigate PLK4 substrates. Similarly, the datasets provide a number of potential PLK4 substrates that could be further investigated. PLK4 is a key chemotherapeutic target and therefore, the work in this thesis can be used to further understand the global effects on the cell when PLK4 activity is inhibited.

8. References

Adon, A. M., Zeng, X., Harrison, M. K., Sannem, S., Kiyokawa, H., Kaldis, P., & Saavedra, H. I. (2010). Cdk2 and Cdk4 regulate the centrosome cycle and are critical mediators of centrosome amplification in p53-null cells. *Mol Cell Biol, 30*(3), 694-710.

Alpert, A. J. (2008). Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. *Anal Chem, 80*(1), 62-76.

Anastassiadis, T., Deacon, S. W., Devarajan, K., Ma, H., & Peterson, J. R. (2011). Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. *Nat Biotechnol, 29*(11), 1039-1045.

Aoki, T., Ueda, S., Kataoka, T., & Satoh, T. (2009). Regulation of mitotic spindle formation by the RhoA guanine nucleotide exchange factor ARHGEF10. *BMC Cell Biol, 10*, 56.

Arquint, C., Gabryjonczyk, A. M., Imseng, S., Bohm, R., Sauer, E., Hiller, S., Nigg, E. A., & Maier, T. (2015). STIL binding to Polo-box 3 of PLK4 regulates centriole duplication. *Elife, 4*.

Arquint, C., & Nigg, E. A. (2016). The PLK4-STIL-SAS-6 module at the core of centriole duplication. *Biochem Soc Trans*, *44*(5), 1253-1263.

Arquint, C., Sonnen, K. F., Stierhof, Y. D., & Nigg, E. A. (2012). Cell-cycle-regulated expression of STIL controls centriole number in human cells. *J Cell Sci, 125*(Pt 5), 1342-1352.

Asteriti, I. A., De Mattia, F., & Guarguaglini, G. (2015). Cross-Talk between AURKA and Plk1 in Mitotic Entry and Spindle Assembly. *Front Oncol, 5*, 283.

Bahtz, R., Seidler, J., Arnold, M., Haselmann-Weiss, U., Antony, C., Lehmann, W. D., & Hoffmann, I. (2012). GCP6 is a substrate of Plk4 and required for centriole duplication. *J Cell Sci, 125*(Pt 2), 486-496.

Bai, M., Ni, J., Wu, J., Wang, B., Shen, S., & Yu, L. (2014). A novel mechanism for activation of Aurora-A kinase by Ajuba. *Gene, 543*(1), 133-139.

Bailey, C. M., Sweet, S. M., Cunningham, D. L., Zeller, M., Heath, J. K., & Cooper, H. J. (2009). SLoMo: automated site localization of modifications from ETD/ECD mass spectra. *J Proteome Res, 8*(4), 1965-1971.

Bailey, F. P., Andreev, V. I., & Eyers, P. A. (2014). The resistance tetrad: amino acid hotspots for kinome-wide exploitation of drug-resistant protein kinase alleles. *Methods Enzymol, 548*, 117-146.

Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C. J., McLauchlan, H., Klevernic, I., Arthur, J. S., Alessi, D. R., & Cohen, P. (2007). The selectivity of protein kinase inhibitors: a further update. *Biochem J, 408*(3), 297-315.

Baldin, V., Lukas, J., Marcote, M. J., Pagano, M., & Draetta, G. (1993). Cyclin D1 is a nuclear protein required for cell cycle progression in G1. *Genes Dev, 7*(5), 812-821.

Barros, T. P., Kinoshita, K., Hyman, A. A., & Raff, J. W. (2005). Aurora A activates D-TACC-Msps complexes exclusively at centrosomes to stabilize centrosomal microtubules. *J Cell Biol, 170*(7), 1039-1046.

Basant, A., Lekomtsev, S., Tse, Y. C., Zhang, D., Longhini, K. M., Petronczki, M., & Glotzer, M. (2015). Aurora B kinase promotes cytokinesis by inducing centralspindlin oligomers that associate with the plasma membrane. *Dev Cell, 33*(2), 204-215.

Basto, R., Brunk, K., Vinadogrova, T., Peel, N., Franz, A., Khodjakov, A., & Raff, J. W. (2008). Centrosome amplification can initiate tumorigenesis in flies. *Cell, 133*(6), 1032-1042.

Batth, T. S., Francavilla, C., & Olsen, J. V. (2014). Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics. *J Proteome Res, 13*(12), 6176-6186.

Bauer, M., Cubizolles, F., Schmidt, A., & Nigg, E. A. (2016). Quantitative analysis of human centrosome architecture by targeted proteomics and fluorescence imaging. *Embo j, 35*(19), 2152-2166.

Bayliss, R., Sardon, T., Vernos, I., & Conti, E. (2003). Structural basis of Aurora-A activation by TPX2 at the mitotic spindle. *Mol Cell, 12*(4), 851-862.

Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E., Villen, J., Li, J., Cohn, M. A., Cantley, L. C., & Gygi, S. P. (2004). Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci U S A, 101*(33), 12130-12135.

Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J., & Gygi, S. P. (2006). A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat Biotechnol, 24*(10), 1285-1292.

Bettencourt-Dias, M., Rodrigues-Martins, A., Carpenter, L., Riparbelli, M., Lehmann, L., Gatt, M. K., Carmo, N., Balloux, F., Callaini, G., & Glover, D. M. (2005). SAK/PLK4 is required for centriole duplication and flagella development. *Curr Biol, 15*(24), 2199-2207.

Beynon, R. J., Doherty, M. K., Pratt, J. M., & Gaskell, S. J. (2005). Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nat Methods, 2*(8), 587-589.

Bischoff, J. R., Anderson, L., Zhu, Y., Mossie, K., Ng, L., Souza, B., Schryver, B., Flanagan, P., Clairvoyant, F., Ginther, C., Chan, C. S., Novotny, M., Slamon, D. J., & Plowman, G. D. (1998). A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. *Embo j, 17*(11), 3052-3065.

Bishop, A. C., Ubersax, J. A., Petsch, D. T., Matheos, D. P., Gray, N. S., Blethrow, J., Shimizu, E., Tsien, J. Z., Schultz, P. G., Rose, M. D., Wood, J. L., Morgan, D. O., & Shokat, K. M. (2000). A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature, 407*(6802), 395-401.

Blagoev, B., & Mann, M. (2006). Quantitative proteomics to study mitogen-activated protein kinases. *Methods, 40*(3), 243-250.

Blay, J. Y., & von Mehren, M. (2011). Nilotinib: a novel, selective tyrosine kinase inhibitor. *Semin Oncol, 38 Suppl 1*, S3-9.

Blethrow, J. D., Glavy, J. S., Morgan, D. O., & Shokat, K. M. (2008). Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates. *Proc Natl Acad Sci U S A, 105*(5), 1442-1447.

Blomberg, I., & Hoffmann, I. (1999). Ectopic expression of Cdc25A accelerates the G(1)/S transition and leads to premature activation of cyclin E- and cyclin A-dependent kinases. *Mol Cell Biol, 19*(9), 6183-6194.

Boersema, P. J., Mohammed, S., & Heck, A. J. (2009a). Phosphopeptide fragmentation and analysis by mass spectrometry. *J Mass Spectrom, 44*(6), 861-878.

Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., & Heck, A. J. (2009b). Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc, 4*(4), 484-494.

Bolton, M. A., Lan, W., Powers, S. E., McCleland, M. L., Kuang, J., & Stukenberg, P. T. (2002). Aurora B kinase exists in a complex with survivin and INCENP and its kinase activity is stimulated by survivin binding and phosphorylation. *Mol Biol Cell, 13*(9), 3064-3077.

Bondarenko, P. V., Chelius, D., & Shaler, T. A. (2002). Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. *Anal Chem, 74*(18), 4741-4749.

Bononi, A., Agnoletto, C., De Marchi, E., Marchi, S., Patergnani, S., Bonora, M., Giorgi, C., Missiroli, S., Poletti, F., Rimessi, A., & Pinton, P. (2011). Protein kinases and phosphatases in the control of cell fate. *Enzyme Res, 2011*, 329098.

Brownlee, C. W., Klebba, J. E., Buster, D. W., & Rogers, G. C. (2011). The Protein Phosphatase 2A regulatory subunit Twins stabilizes Plk4 to induce centriole amplification. *J Cell Biol, 195*(2), 231-243.

Budhu, A. S., & Wang, X. W. (2005). Loading and unloading: orchestrating centrosome duplication and spindle assembly by Ran/Crm1. *Cell Cycle, 4*(11), 1510-1514.

Burkard, M. E., Randall, C. L., Larochelle, S., Zhang, C., Shokat, K. M., Fisher, R. P., & Jallepalli, P. V. (2007). Chemical genetics reveals the requirement for Polo-like kinase 1 activity in positioning RhoA and triggering cytokinesis in human cells. *Proc Natl Acad Sci U S A, 104*(11), 4383-4388.

Burns, T. F., Fei, P., Scata, K. A., Dicker, D. T., & El-Deiry, W. S. (2003). Silencing of the novel p53 target gene Snk/Plk2 leads to mitotic catastrophe in paclitaxel (taxol)-exposed cells. *Mol Cell Biol, 23*(16), 5556-5571.

Bury, L., Coelho, P. A., Simeone, A., Ferries, S., Eyers, C. E., Eyers, P. A., Zernicka-Goetz, M., & Glover, D. M. (2017). Plk4 and Aurora A cooperate in the initiation of acentriolar spindle assembly in mammalian oocytes. *J Cell Biol, 216*(11), 3571-3590.

Butch, E. R., & Guan, K. L. (1996). Characterization of ERK1 activation site mutants and the effect on recognition by MEK1 and MEK2. *J Biol Chem, 271*(8), 4230-4235.

Cairns, R. A., Harris, I. S., & Mak, T. W. (2011). Regulation of cancer cell metabolism. *Nat Rev Cancer, 11*(2), 85-95.

Cajanek, L., Glatter, T., & Nigg, E. A. (2015). The E3 ubiquitin ligase Mib1 regulates Plk4 and centriole biogenesis. *J Cell Sci, 128*(9), 1674-1682.

Casolaro, A., Golay, J., Albanese, C., Ceruti, R., Patton, V., Cribioli, S., Pezzoni, A., Losa, M., Texido, G., Giussani, U., Marchesi, F., Amboldi, N., Valsasina, B.,

Bungaro, S., Cazzaniga, G., Rambaldi, A., Introna, M., Pesenti, E., & Alzani, R. (2013). The Polo-Like Kinase 1 (PLK1) inhibitor NMS-P937 is effective in a new model of disseminated primary CD56+ acute monoblastic leukaemia. *PLoS One, 8*(3), e58424.

Cazales, M., Schmitt, E., Montembault, E., Dozier, C., Prigent, C., & Ducommun, B. (2005). CDC25B phosphorylation by Aurora-A occurs at the G2/M transition and is inhibited by DNA damage. *Cell Cycle, 4*(9), 1233-1238.

Chalkley, R. J., & Clauser, K. R. (2012). Modification site localization scoring: strategies and performance. *Mol Cell Proteomics, 11*(5), 3-14.

Chapman, J. D., Goodlett, D. R., & Masselon, C. D. (2014). Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. *Mass Spectrom Rev, 33*(6), 452-470.

Cheeseman, I. M., Chappie, J. S., Wilson-Kubalek, E. M., & Desai, A. (2006). The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. *Cell, 127*(5), 983-997.

Cheng, M., Olivier, P., Diehl, J. A., Fero, M., Roussel, M. F., Roberts, J. M., & Sherr, C. J. (1999). The p21(Cip1) and p27(Kip1) CDK 'inhibitors' are essential activators of cyclin D-dependent kinases in murine fibroblasts. *Embo j, 18*(6), 1571-1583.

Chou, M. F., & Schwartz, D. (2011). Biological sequence motif discovery using motif-x. *Curr Protoc Bioinformatics, Chapter 13*, Unit 13.15-24.

Choudhary, C., & Mann, M. (2010). Decoding signalling networks by mass spectrometry-based proteomics. *Nat Rev Mol Cell Biol, 11*(6), 427-439.

Cizmecioglu, O., Krause, A., Bahtz, R., Ehret, L., Malek, N., & Hoffmann, I. (2012). Plk2 regulates centriole duplication through phosphorylation-mediated degradation of Fbxw7 (human Cdc4). *J Cell Sci, 125*(Pt 4), 981-992.

Coelho, P. A., Bury, L., Shahbazi, M. N., Liakath-Ali, K., Tate, P. H., Wormald, S., Hindley, C. J., Huch, M., Archer, J., Skarnes, W. C., Zernicka-Goetz, M., & Glover, D. M. (2015). Over-expression of Plk4 induces centrosome amplification, loss of primary cilia and associated tissue hyperplasia in the mouse. *Open Biol, 5*(12), 150209.

Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., & Gevaert, K. (2009). Improved visualization of protein consensus sequences by iceLogo. *Nat Methods, 6*(11), 786-787.

Coverley, D., Laman, H., & Laskey, R. A. (2002). Distinct roles for cyclins E and A during DNA replication complex assembly and activation. *Nat Cell Biol, 4*(7), 523-528.

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol, 26*(12), 1367-1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., & Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res, 10*(4), 1794-1805.

Cox, K. A., Cleven, C. D., & Cooks, R. G. (1995). Mass shifts and local space charge effects observed in the quadrupole ion trap at higher resolution. *International Journal of Mass Spectrometry and Ion Processes, 144*(1), 47-65.

Cunha-Ferreira, I., Bento, I., Pimenta-Marques, A., Jana, S. C., Lince-Faria, M., Duarte, P., Borrego-Pinto, J., Gilberto, S., Amado, T., Brito, D., Rodrigues-Martins, A., Debski, J., Dzhindzhev, N., & Bettencourt-Dias, M. (2013). Regulation of autophosphorylation controls PLK4 self-destruction and centriole number. *Curr Biol, 23*(22), 2245-2254.

Dammermann, A., Muller-Reichert, T., Pelletier, L., Habermann, B., Desai, A., & Oegema, K. (2004). Centriole assembly requires both centriolar and pericentriolar material proteins. *Dev Cell, 7*(6), 815-829.

Dar, A. C., & Shokat, K. M. (2011). The evolution of protein kinase inhibitors from antagonists to agonists of cellular signaling. *Annu Rev Biochem, 80*, 769-795.

de Carcer, G., Escobar, B., Higuero, A. M., Garcia, L., Anson, A., Perez, G., Mollejo, M., Manning, G., Melendez, B., Abad-Rodriguez, J., & Malumbres, M. (2011). Plk5, a polo box domain-only protein with specific roles in neuron differentiation and glioblastoma suppression. *Mol Cell Biol, 31*(6), 1225-1239.

de Carcer, G., Venkateswaran, S. V., Salgueiro, L., El Bakkali, A., Somogyi, K., Rowald, K., Montanes, P., Sanclemente, M., Escobar, B., de Martino, A., McGranahan, N., Malumbres, M., & Sotillo, R. (2018). Plk1 overexpression induces chromosomal instability and suppresses tumor development. *Nat Commun, 9*(1), 3012.

DeGnore, J. P., & Qin, J. (1998). Fragmentation of phosphopeptides in an ion trap mass spectrometer. *J Am Soc Mass Spectrom, 9*(11), 1175-1188.

Deininger, M. W., Goldman, J. M., Lydon, N., & Melo, J. V. (1997). The tyrosine kinase inhibitor CGP57148B selectively inhibits the growth of BCR-ABL-positive cells. *Blood, 90*(9), 3691-3698.

Delmotte, N., Lasaosa, M., Tholey, A., Heinzle, E., & Huber, C. G. (2007). Two-dimensional reversed-phase x ion-pair reversed-phase HPLC: an alternative approach to high-resolution peptide separation for shotgun proteome analysis. *J Proteome Res, 6*(11), 4363-4373.

DeLuca, K. F., Meppelink, A., Broad, A. J., Mick, J. E., Peersen, O. B., Pektas, S., Lens, S. M. A., & DeLuca, J. G. (2018). Aurora A kinase phosphorylates Hec1 to regulate metaphase kinetochore-microtubule dynamics. *J Cell Biol, 217*(1), 163-177.

Di, W., Khan, M., Rasul, A., Sun, M., Sui, Y., Zhong, L., Yang, L., Zhu, Q., Feng, L., & Ma, T. (2014). Isoalantolactone inhibits constitutive NF-kappaB activation and induces reactive oxygen species-mediated apoptosis in osteosarcoma U2OS cells through mitochondrial dysfunction. *Oncol Rep, 32*(4), 1585-1593.

Diedrich, J. K., Pinto, A. F., & Yates, J. R., 3rd. (2013). Energy dependence of HCD on peptide fragmentation: stepped collisional energy finds the sweet spot. *J Am Soc Mass Spectrom, 24*(11), 1690-1699.

Doherty, M. K., Hammond, D. E., Clague, M. J., Gaskell, S. J., & Beynon, R. J. (2009). Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. *J Proteome Res, 8*(1), 104-112.

Dole, M., L. Mack, L., L. Hines, R., Mobley, R. C., Ferguson, L. D., & Alice, M. B. (1968). *Molecular Beams of Macroions* (Vol. 49).

Dongré, A. R., Jones, J. L., Somogyi, Á., & Wysocki, V. H. (1996). Influence of Peptide Composition, Gas-Phase Basicity, and Chemical Modification on

Fragmentation Efficiency: Evidence for the Mobile Proton Model. *Journal of the American Chemical Society, 118*(35), 8365-8374.

Dorfer, V., Pichler, P., Stranzl, T., Stadlmann, J., Taus, T., Winkler, S., & Mechtler, K. (2014). MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. *J Proteome Res, 13*(8), 3679-3684.

Douglas, D. J., Frank, A. J., & Mao, D. (2005). Linear ion traps in mass spectrometry. *Mass Spectrom Rev, 24*(1), 1-29.

Doxsey, S. J., Stein, P., Evans, L., Calarco, P. D., & Kirschner, M. (1994). Pericentrin, a highly conserved centrosome protein involved in microtubule organization. *Cell, 76*(4), 639-650.

Dutertre, S., Descamps, S., & Prigent, C. (2002). On the role of aurora-A in centrosome function. *Oncogene, 21*(40), 6175-6183.

Dzhindzhev, N. S., Tzolovsky, G., Lipinszki, Z., Abdelaziz, M., Debski, J., Dadlez, M., & Glover, D. M. (2017). Two-step phosphorylation of Ana2 by Plk4 is required for the sequential loading of Ana2 and Sas6 to initiate procentriole formation. *Open Biol, 7*(12).

Dzhindzhev, N. S., Tzolovsky, G., Lipinszki, Z., Schneider, S., Lattao, R., Fu, J., Debski, J., Dadlez, M., & Glover, D. M. (2014). Plk4 phosphorylates Ana2 to trigger Sas6 recruitment and procentriole formation. *Curr Biol, 24*(21), 2526-2532.

Elias, J. E., & Gygi, S. P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods, 4*(3), 207-214.

Eliuk, S., & Makarov, A. (2015). Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annu Rev Anal Chem (Palo Alto Calif), 8*, 61-80.

Emanuel, S., Rugg, C. A., Gruninger, R. H., Lin, R., Fuentes-Pesquera, A., Connolly, P. J., Wetter, S. K., Hollister, B., Kruger, W. W., Napier, C., Jolliffe, L., & Middleton, S. A. (2005). The in vitro and in vivo effects of JNJ-7706621: a dual inhibitor of cyclin-dependent kinases and aurora kinases. *Cancer Res, 65*(19), 9038-9046.

Enami, M., & Ishihama, A. (1984). Protein phosphorylation in Escherichia coli and purification of a protein kinase. *J Biol Chem, 259*(1), 526-533.

Endicott, J. A., Noble, M. E., & Johnson, L. N. (2012). The structural basis for control of eukaryotic protein kinases. *Annu Rev Biochem, 81*, 587-613.

Eng, J. K., McCormack, A. L., & Yates, J. R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom, 5*(11), 976-989.

Espadas, G., Borras, E., Chiva, C., & Sabido, E. (2017). Evaluation of different peptide fragmentation types and mass analyzers in data-dependent methods using an Orbitrap Fusion Lumos Tribrid mass spectrometer. *Proteomics, 17*(9).

Eyers, P. A., Churchill, M. E., & Maller, J. L. (2005). The Aurora A and Aurora B protein kinases: a single amino acid difference controls intrinsic activity and activation by TPX2. *Cell Cycle, 4*(6), 784-789.

Eyers, P. A., Erikson, E., Chen, L. G., & Maller, J. L. (2003). A novel mechanism for activation of the protein kinase Aurora A. *Curr Biol, 13*(8), 691-697.

Fabian, M. A., Biggs, W. H., 3rd, Treiber, D. K., Atteridge, C. E., Azimioara, M. D., Benedetti, M. G., Carter, T. A., Ciceri, P., Edeen, P. T., Floyd, M., Ford, J. M., Galvin, M., Gerlach, J. L., Grotzfeld, R. M., Herrgard, S., Insko, D. E., Insko, M.

A., Lai, A. G., Lelias, J. M., Mehta, S. A., Milanov, Z. V., Velasco, A. M., Wodicka, L. M., Patel, H. K., Zarrinkar, P. P., & Lockhart, D. J. (2005). A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat Biotechnol, 23*(3), 329-336.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whitehouse, C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science, 246*(4926), 64-71.

Ferries, S., Perkins, S., Brownridge, P. J., Campbell, A., Eyers, P. A., Jones, A. R. & Eyers, C. E. (2017). 'Evaluation of Parameters for Confident Phosphorylation Site Localization Using an Orbitrap Fusion Tribrid Mass Spectrometer'. *J Proteome Res*. *16*. 3448-3459.

Fermin, D., Walmsley, S. J., Gingras, A. C., Choi, H., & Nesvizhskii, A. I. (2013). LuciPHOr: algorithm for phosphorylation site localization with false localization rate estimation using modified target-decoy approach. *Mol Cell Proteomics, 12*(11), 3409-3419.

Firat-Karalar, E. N., Rauniyar, N., Yates, J. R., 3rd, & Stearns, T. (2014). Proximity interactions among centrosome components identify regulators of centriole duplication. *Curr Biol, 24*(6), 664-670.

Fleitz, A., Nieves, E., Madrid-Aliste, C., Fentress, S. J., Sibley, L. D., Weiss, L. M., Angeletti, R. H., & Che, F. Y. (2013). Enhanced detection of multiply phosphorylated peptides and identification of their sites of modification. *Anal Chem, 85*(18), 8566-8576.

Fode, C., Motro, B., Yousefi, S., Heffernan, M., & Dennis, J. W. (1994). Sak, a murine protein-serine/threonine kinase that is related to the Drosophila polo kinase and involved in cell proliferation. *Proc Natl Acad Sci U S A, 91*(14), 6388-6392.

Fournier, M., Orpinell, M., Grauffel, C., Scheer, E., Garnier, J. M., Ye, T., Chavant, V., Joint, M., Esashi, F., Dejaegere, A., Gonczy, P., & Tora, L. (2016). KAT2A/KAT2B-targeted acetylome reveals a role for PLK4 acetylation in preventing centrosome amplification. *Nat Commun, 7*, 13227.

Frese, C. K., Altelaar, A. F., van den Toorn, H., Nolting, D., Griep-Raming, J., Heck, A. J., & Mohammed, S. (2012). Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry. *Anal Chem, 84*(22), 9668-9673.

Frese, C. K., Zhou, H., Taus, T., Altelaar, A. F., Mechtler, K., Heck, A. J., & Mohammed, S. (2013). Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (EThcD). *J Proteome Res, 12*(3), 1520-1525.

Gallini, S., Carminati, M., De Mattia, F., Pirovano, L., Martini, E., Oldani, A., Asteriti, I. A., Guarguaglini, G., & Mapelli, M. (2016). NuMA Phosphorylation by Aurora-A Orchestrates Spindle Orientation. *Curr Biol, 26*(4), 458-469.

Geiger, T., Wisniewski, J. R., Cox, J., Zanivan, S., Kruger, M., Ishihama, Y., & Mann, M. (2011). Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. *Nat Protoc, 6*(2), 147-157.

Giet, R., & Glover, D. M. (2001). Drosophila aurora B kinase is required for histone H3 phosphorylation and condensin recruitment during chromosome

condensation and to organize the central spindle during cytokinesis. *J Cell Biol, 152*(4), 669-682.

Giet, R., & Prigent, C. (1999). Aurora/Ipl1p-related kinases, a new oncogenic family of mitotic serine-threonine kinases. *J Cell Sci, 112 (Pt 21)*, 3591-3601.

Gilar, M., Olivova, P., Daly, A. E., & Gebler, J. C. (2005). Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *J Sep Sci, 28*(14), 1694-1703.

Gilmartin, A. G., Bleam, M. R., Richter, M. C., Erskine, S. G., Kruger, R. G., Madden, L., Hassler, D. F., Smith, G. K., Gontarek, R. R., Courtney, M. P., Sutton, D., Diamond, M. A., Jackson, J. R., & Laquerre, S. G. (2009). Distinct concentration-dependent effects of the polo-like kinase 1-specific inhibitor GSK461364A, including differential effect on apoptosis. *Cancer Res, 69*(17), 6969-6977.

Glotzer, M. (2005). The molecular requirements for cytokinesis. *Science, 307*(5716), 1735-1739.

Glover, D. M., Leibowitz, M. H., McLean, D. A., & Parry, H. (1995). Mutations in aurora prevent centrosome separation leading to the formation of monopolar spindles. *Cell, 81*(1), 95-105.

Goeringer, D. E., Whitten, W. B., Ramsey, J. M., McLuckey, S. A., & Glish, G. L. (1992). Theory of high-resolution mass spectrometry achieved via resonance ejection in the quadrupole ion trap. *Anal Chem, 64*(13), 1434-1439.

Goetz, S. C., & Anderson, K. V. (2010). The primary cilium: a signalling centre during vertebrate development. *Nat Rev Genet, 11*(5), 331-344.

Gomez-Ferreria, M. A., Rath, U., Buster, D. W., Chanda, S. K., Caldwell, J. S., Rines, D. R., & Sharp, D. J. (2007). Human Cep192 is required for mitotic centrosome and spindle assembly. *Curr Biol, 17*(22), 1960-1966.

Gonzalez-Loyola, A., Fernandez-Miranda, G., Trakala, M., Partida, D., Samejima, K., Ogawa, H., Canamero, M., de Martino, A., Martinez-Ramirez, A., de Carcer, G., Perez de Castro, I., Earnshaw, W. C., & Malumbres, M. (2015). Aurora B Overexpression Causes Aneuploidy and p21Cip1 Repression during Tumor Development. *Mol Cell Biol, 35*(20), 3566-3578.

Good, D. M., Wirtala, M., McAlister, G. C., & Coon, J. J. (2007). Performance characteristics of electron transfer dissociation mass spectrometry. *Mol Cell Proteomics, 6*(11), 1942-1951.

Gorgun, G., Calabrese, E., Hideshima, T., Ecsedy, J., Perrone, G., Mani, M., Ikeda, H., Bianchi, G., Hu, Y., Cirstea, D., Santo, L., Tai, Y. T., Nahar, S., Zheng, M., Bandi, M., Carrasco, R. D., Raje, N., Munshi, N., Richardson, P., & Anderson, K. C. (2010). A novel Aurora-A kinase inhibitor MLN8237 induces cytotoxicity and cell-cycle arrest in multiple myeloma. *Blood, 115*(25), 5202-5213.

Gorre, M. E., Mohammed, M., Ellwood, K., Hsu, N., Paquette, R., Rao, P. N., & Sawyers, C. L. (2001). Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science, 293*(5531), 876-880.

Grosstessner-Hain, K., Hegemann, B., Novatchkova, M., Rameseder, J., Joughin, B. A., Hudecz, O., Roitinger, E., Pichler, P., Kraut, N., Yaffe, M. B., Peters, J. M., & Mechtler, K. (2011). Quantitative phospho-proteomics to investigate the polo-like kinase 1-dependent phospho-proteome. *Mol Cell Proteomics, 10*(11), M111.008540.

Guderian, G., Westendorf, J., Uldschmid, A., & Nigg, E. A. (2010). Plk4 trans-autophosphorylation regulates centriole number by controlling betaTrCP-mediated degradation. *J Cell Sci, 123*(Pt 13), 2163-2169.

Habedanck, R., Stierhof, Y. D., Wilkinson, C. J., & Nigg, E. A. (2005). The Polo kinase Plk4 functions in centriole duplication. *Nat Cell Biol, 7*(11), 1140-1146.

Han, S., Kim, S., Bahl, S., Li, L., Burande, C. F., Smith, N., James, M., Beauchamp, R. L., Bhide, P., DiAntonio, A., & Ramesh, V. (2012). The E3 ubiquitin ligase protein associated with Myc (Pam) regulates mammalian/mechanistic target of rapamycin complex 1 (mTORC1) signaling in vivo through N- and C-terminal domains. *J Biol Chem, 287*(36), 30063-30072.

Han, X., Aslanian, A., & Yates, J. R., 3rd. (2008). Mass spectrometry for proteomics. *Curr Opin Chem Biol, 12*(5), 483-490.

Hanks, S. K., & Hunter, T. (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *Faseb j, 9*(8), 576-596.

Hanks, S. K., Quinn, A. M., & Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science, 241*(4861), 42-52.

Hannak, E., Kirkham, M., Hyman, A. A., & Oegema, K. (2001). Aurora-A kinase is required for centrosome maturation in Caenorhabditis elegans. *J Cell Biol, 155*(7), 1109-1116.

Hardman, G., Perkins, S., Ruan, Z., Kannan, N., Brownridge, P., Byrne, D. P., Eyers, P. A., Jones, A. R., & Eyers, C. E. (2017). Extensive non-canonical phosphorylation in human cells revealed using strong-anion exchange-mediated phosphoproteomics. *bioRxiv*.

Harrington, E. A., Bebbington, D., Moore, J., Rasmussen, R. K., Ajose-Adeogun, A. O., Nakayama, T., Graham, J. A., Demur, C., Hercend, T., Diu-Hercend, A., Su, M., Golec, J. M. C., & Miller, K. M. (2004). VX-680, a potent and selective small-molecule inhibitor of the Aurora kinases, suppresses tumor growth in vivo. *Nature Medicine, 10*, 262.

Hatch, E. M., Kulukian, A., Holland, A. J., Cleveland, D. W., & Stearns, T. (2010). Cep152 interacts with Plk4 and is required for centriole duplication. *J Cell Biol, 191*(4), 721-729.

Hauf, S., Roitinger, E., Koch, B., Dittrich, C. M., Mechtler, K., & Peters, J. M. (2005). Dissociation of cohesin from chromosome arms and loss of arm cohesion during early mitosis depends on phosphorylation of SA2. *PLoS Biol, 3*(3), e69.

Hebert, A. S., Richards, A. L., Bailey, D. J., Ulbrich, A., Coughlin, E. E., Westphall, M. S., & Coon, J. J. (2014). The one hour yeast proteome. *Mol Cell Proteomics, 13*(1), 339-347.

Hemmer, W., McGlone, M., Tsigelny, I., & Taylor, S. S. (1997). Role of the glycine triad in the ATP-binding site of cAMP-dependent protein kinase. *J Biol Chem, 272*(27), 16946-16954.

Hirota, T., Kunitoku, N., Sasayama, T., Marumoto, T., Zhang, D., Nitta, M., Hatakeyama, K., & Saya, H. (2003). Aurora-A and an interacting activator, the LIM protein Ajuba, are required for mitotic commitment in human cells. *Cell, 114*(5), 585-598.

Holland, A. J., Fachinetti, D., Zhu, Q., Bauer, M., Verma, I. M., Nigg, E. A., & Cleveland, D. W. (2012). The autoregulated instability of Polo-like kinase 4 limits centrosome duplication to once per cell cycle. *Genes Dev, 26*(24), 2684-2689.

Holland, A. J., Lan, W., Niessen, S., Hoover, H., & Cleveland, D. W. (2010). Polo-like kinase 4 kinase activity limits centrosome overduplication by autoregulating its own stability. *J Cell Biol, 188*(2), 191-198.

Honda, R., Korner, R., & Nigg, E. A. (2003). Exploring the functional interactions between Aurora B, INCENP, and survivin in mitosis. *Mol Biol Cell, 14*(8), 3325-3341.

Hori, A., Barnouin, K., Snijders, A. P., & Toda, T. (2016). A non-canonical function of Plk4 in centriolar satellite integrity and ciliogenesis through PCM1 phosphorylation. *EMBO Rep, 17*(3), 326-337.

Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., & Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res, 43*(Database issue), D512-520.

Hsu, J. L., Huang, S. Y., Chow, N. H., & Chen, S. H. (2003). Stable-isotope dimethyl labeling for quantitative proteomics. *Anal Chem, 75*(24), 6843-6852.

Huang da, W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc, 4*(1), 44-57.

Hunt, D. F., Yates, J. R., 3rd, Shabanowitz, J., Winston, S., & Hauer, C. R. (1986). Protein sequencing by tandem mass spectrometry. *Proc Natl Acad Sci U S A, 83*(17), 6233-6237.

Hunter, T. (1995). Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell, 80*(2), 225-236.

Huse, M., & Kuriyan, J. (2002). The conformational plasticity of protein kinases. *Cell, 109*(3), 275-282.

Hutterer, A., Berdnik, D., Wirtz-Peitz, F., Zigman, M., Schleiffer, A., & Knoblich, J. A. (2006). Mitotic activation of the kinase Aurora-A requires its binding partner Bora. *Dev Cell, 11*(2), 147-157.

Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., & Dunker, A. K. (2004). The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res, 32*(3), 1037-1049.

Ikeda, M. A., Jakoi, L., & Nevins, J. R. (1996). A unique role for the Rb protein in controlling E2F accumulation during cell growth and differentiation. *Proc Natl Acad Sci U S A, 93*(8), 3215-3220.

Ishida, S., Huang, E., Zuzan, H., Spang, R., Leone, G., West, M., & Nevins, J. R. (2001). Role for E2F in control of both DNA replication and mitotic functions as revealed from DNA microarray analysis. *Mol Cell Biol, 21*(14), 4684-4699.

Ishihama, Y., Wei, F. Y., Aoshima, K., Sato, T., Kuromitsu, J., & Oda, Y. (2007). Enhancement of the efficiency of phosphoproteomic identification by removing phosphates after phosphopeptide enrichment. *J Proteome Res, 6*(3), 1139-1144.

Jackman, M., Lindon, C., Nigg, E. A., & Pines, J. (2003). Active cyclin B1-Cdk1 first appears on centrosomes in prophase. *Nat Cell Biol, 5*(2), 143-148.

James, M. K., Ray, A., Leznova, D., & Blain, S. W. (2008). Differential modification of p27Kip1 controls its cyclin D-cdk4 inhibitory activity. *Mol Cell Biol, 28*(1), 498-510.

Jedrychowski, M. P., Huttlin, E. L., Haas, W., Sowa, M. E., Rad, R., & Gygi, S. P. (2011). Evaluation of HCD- and CID-type fragmentation within their respective detection platforms for murine phosphoproteomics. *Mol Cell Proteomics, 10*(12), M111.009910.

Johnson, E. F., Stewart, K. D., Woods, K. W., Giranda, V. L., & Luo, Y. (2007). Pharmacological and functional comparison of the polo-like kinase family: insight into inhibitor and substrate specificity. *Biochemistry, 46*(33), 9551-9563.

Johnson, L. N., Noble, M. E., & Owen, D. J. (1996). Active and inactive protein kinases: structural basis for regulation. *Cell, 85*(2), 149-158.

Joukov, V., De Nicolo, A., Rodriguez, A., Walter, J. C., & Livingston, D. M. (2010). Centrosomal protein of 192 kDa (Cep192) promotes centrosome-driven spindle assembly by engaging in organelle-specific Aurora A activation. *Proc Natl Acad Sci U S A, 107*(49), 21022-21027.

Joukov, V., Walter, J. C., & De Nicolo, A. (2014). The Cep192-organized aurora A-Plk1 cascade is essential for centrosome cycle and bipolar spindle assembly. *Mol Cell, 55*(4), 578-591.

Kammers, K., Cole, R. N., Tiengwe, C., & Ruczinski, I. (2015). Detecting Significant Changes in Protein Abundance. *EuPA Open Proteom, 7*, 11-19.

Kato, J. Y., Matsuoka, M., Strom, D. K., & Sherr, C. J. (1994). Regulation of cyclin D-dependent kinase 4 (cdk4) by cdk4-activating kinase. *Mol Cell Biol, 14*(4), 2713-2721.

Kazazian, K., Go, C., Wu, H., Brashavitskaya, O., Xu, R., Dennis, J. W., Gingras, A. C., & Swallow, C. J. (2017). Plk4 Promotes Cancer Invasion and Metastasis through Arp2/3 Complex Regulation of the Actin Cytoskeleton. *Cancer Res, 77*(2), 434-447.

Kebarle, P., & Verkerk, U. H. (2009). Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom Rev, 28*(6), 898-917.

Kelstrup, C. D., Bekker-Jensen, D. B., Arrey, T. N., Hogrebe, A., Harder, A., & Olsen, J. V. (2018). Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics. *J Proteome Res, 17*(1), 727-738.

Kelstrup, C. D., Jersie-Christensen, R. R., Batth, T. S., Arrey, T. N., Kuehn, A., Kellmann, M., & Olsen, J. V. (2014). Rapid and deep proteomes by faster sequencing on a benchtop quadrupole ultra-high-field Orbitrap mass spectrometer. *J Proteome Res, 13*(12), 6187-6195.

Kettenbach, A. N., Rush, J., & Gerber, S. A. (2011a). Absolute quantification of protein and post-translational modification abundance with stable isotope-labeled synthetic peptides. *Nat Protoc, 6*(2), 175-186.

Kettenbach, A. N., Schweppe, D. K., Faherty, B. K., Pechenick, D., Pletnev, A. A., & Gerber, S. A. (2011b). Quantitative phosphoproteomics identifies substrates and functional modules of Aurora and Polo-like kinase activities in mitotic cells. *Sci Signal, 4*(179), rs5.

Kim, J., Camp, D. G., 2nd, & Smith, R. D. (2004). Improved detection of multi-phosphorylated peptides in the presence of phosphoric acid in liquid chromatography/mass spectrometry. *J Mass Spectrom, 39*(2), 208-215.

Kim, M. S., & Pandey, A. (2012). Electron transfer dissociation mass spectrometry in proteomics. *Proteomics, 12*(4-5), 530-542.

Kim, M. S., Zhong, J., Kandasamy, K., Delanghe, B., & Pandey, A. (2011). Systematic evaluation of alternating CID and ETD fragmentation for phosphorylated peptides. *Proteomics, 11*(12), 2568-2572.

Kim, T. S., Park, J. E., Shukla, A., Choi, S., Murugan, R. N., Lee, J. H., Ahn, M., Rhee, K., Bang, J. K., Kim, B. Y., Loncarek, J., Erikson, R. L., & Lee, K. S. (2013). Hierarchical recruitment of Plk4 and regulation of centriole biogenesis by two centrosomal scaffolds, Cep192 and Cep152. *Proc Natl Acad Sci U S A, 110*(50), E4849-4857.

Kirschner, M., & Mitchison, T. (1986). Beyond self-assembly: from microtubules to morphogenesis. *Cell, 45*(3), 329-342.

Klaeger, S., Heinzlmeir, S., Wilhelm, M., Polzer, H., Vick, B., Koenig, P. A., Reinecke, M., Ruprecht, B., Petzoldt, S., Meng, C., Zecha, J., Reiter, K., Qiao, H., Helm, D., Koch, H., Schoof, M., Canevari, G., Casale, E., Depaolini, S. R., Feuchtinger, A., Wu, Z., Schmidt, T., Rueckert, L., Becker, W., Huenges, J., Garz, A. K., Gohlke, B. O., Zolg, D. P., Kayser, G., Vooder, T., Preissner, R., Hahne, H., Tonisson, N., Kramer, K., Gotze, K., Bassermann, F., Schlegl, J., Ehrlich, H. C., Aiche, S., Walch, A., Greif, P. A., Schneider, S., Felder, E. R., Ruland, J., Medard, G., Jeremias, I., Spiekermann, K., & Kuster, B. (2017). The target landscape of clinical kinase drugs. *Science, 358*(6367).

Klebba, J. E., Buster, D. W., McLamarrah, T. A., Rusan, N. M., & Rogers, G. C. (2015). Autoinhibition and relief mechanism for Polo-like kinase 4. *Proc Natl Acad Sci U S A, 112*(7), E657-666.

Klebba, J. E., Buster, D. W., Nguyen, A. L., Swatkoski, S., Gucek, M., Rusan, N. M., & Rogers, G. C. (2013). Polo-like kinase 4 autodestructs by generating its Slimb-binding phosphodegron. *Curr Biol, 23*(22), 2255-2261.

Kleylein-Sohn, J., Westendorf, J., Le Clech, M., Habedanck, R., Stierhof, Y. D., & Nigg, E. A. (2007). Plk4-induced centriole biogenesis in human cells. *Dev Cell, 13*(2), 190-202.

Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S., & Sowadski, J. M. (1991). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science, 253*(5018), 407-414.

Knudsen, E. S., & Wang, J. Y. (1997). Dual mechanisms for the inhibition of E2F binding to RB by cyclin-dependent kinase-mediated RB phosphorylation. *Mol Cell Biol, 17*(10), 5771-5783.

Ko, M. A., Rosario, C. O., Hudson, J. W., Kulkarni, S., Pollett, A., Dennis, J. W., & Swallow, C. J. (2005). Plk4 haploinsufficiency causes mitotic infidelity and carcinogenesis. *Nat Genet, 37*(8), 883-888.

Kochanski, R. S., & Borisy, G. G. (1990). Mode of centriole duplication and distribution. *J Cell Biol, 110*(5), 1599-1605.

Konermann, L., Ahadi, E., Rodriguez, A. D., & Vahidi, S. (2013). Unraveling the mechanism of electrospray ionization. *Anal Chem, 85*(1), 2-9.

Kornev, A. P., Haste, N. M., Taylor, S. S., & Eyck, L. F. (2006). Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci U S A, 103*(47), 17783-17788.

Kratz, A. S., Barenz, F., Richter, K. T., & Hoffmann, I. (2015). Plk4-dependent phosphorylation of STIL is required for centriole duplication. *Biol Open, 4*(3), 370-377.

Kwon, Y. W., Kim, I. J., Wu, D., Lu, J., Stock, W. A., Jr., Liu, Y., Huang, Y., Kang, H. C., DelRosario, R., Jen, K. Y., Perez-Losada, J., Wei, G., Balmain, A., & Mao, J. H. (2012). Pten regulates Aurora-A and cooperates with Fbxw7 in modulating radiation-induced tumor development. *Mol Cancer Res, 10*(6), 834-844.

LaBaer, J., Garrett, M. D., Stevenson, L. F., Slingerland, J. M., Sandhu, C., Chou, H. S., Fattaey, A., & Harlow, E. (1997). New functional activities for the p21 family of CDK inhibitors. *Genes Dev, 11*(7), 847-862.

Lanucara, F., Lam, C., Mann, J., Monie, T. P., Colombo, S. A., Holman, S. W., Boyd, J., Dange, M. C., Mann, D. A., White, M. R., & Eyers, C. E. (2016). Dynamic phosphorylation of RelA on Ser42 and Ser45 in response to TNFalpha stimulation regulates DNA binding and transcription. *Open Biol, 6*(7).

Lanucara, F., Lee, D. C., & Eyers, C. E. (2014). Unblocking the sink: improved CID-based analysis of phosphorylated peptides by enzymatic removal of the basic C-terminal residue. *J Am Soc Mass Spectrom, 25*(2), 214-225.

Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., & Jorgensen, T. J. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics, 4*(7), 873-886.

Lauper, N., Beck, A. R., Cariou, S., Richman, L., Hofmann, K., Reith, W., Slingerland, J. M., & Amati, B. (1998). Cyclin E2: a novel CDK2 partner in the late G1 and S phases of the mammalian cell cycle. *Oncogene, 17*(20), 2637-2643.

Lee, A. K., & Potts, P. R. (2017). A Comprehensive Guide to the MAGE Family of Ubiquitin Ligases. *J Mol Biol, 429*(8), 1114-1142.

Lee, K., & Rhee, K. (2011). PLK1 phosphorylation of pericentrin initiates centrosome maturation at the onset of mitosis. *J Cell Biol, 195*(7), 1093-1101.

Lee, M., Seo, M. Y., Chang, J., Hwang, D. S., & Rhee, K. (2017). PLK4 phosphorylation of CP110 is required for efficient centriole assembly. *Cell Cycle, 16*(12), 1225-1234.

Leidel, S., Delattre, M., Cerutti, L., Baumer, K., & Gonczy, P. (2005). SAS-6 defines a protein family required for centrosome duplication in C. elegans and in human cells. *Nat Cell Biol, 7*(2), 115-125.

Leung, G. C., Ho, C. S., Blasutig, I. M., Murphy, J. M., & Sicheri, F. (2007). Determination of the Plk4/Sak consensus phosphorylation motif using peptide spots arrays. *FEBS Lett, 581*(1), 77-83.

Leung, G. C., Hudson, J. W., Kozarova, A., Davidson, A., Dennis, J. W., & Sicheri, F. (2002). The Sak polo-box comprises a structural domain sufficient for mitotic subcellular localization. *Nat Struct Biol, 9*(10), 719-724.

Li, Y., Zhang, Z. F., Chen, J., Huang, D., Ding, Y., Tan, M. H., Qian, C. N., Resau, J. H., Kim, H., & Teh, B. T. (2010). VX680/MK-0457, a potent and selective Aurora kinase inhibitor, targets both tumor and endothelial cells in clear cell renal cell carcinoma. *Am J Transl Res, 2*(3), 296-308.

Li, Z., Dai, K., Wang, C., Song, Y., Gu, F., Liu, F., & Fu, L. (2016). Expression of Polo-Like Kinase 4(PLK4) in Breast Cancer and Its Response to Taxane-Based Neoadjuvant Chemotherapy. *J Cancer, 7*(9), 1125-1132.

Lim, N. R., Yeap, Y. Y., Zhao, T. T., Yip, Y. Y., Wong, S. C., Xu, D., Ang, C. S., Williamson, N. A., Xu, Z., Bogoyevitch, M. A., & Ng, D. C. (2015). Opposing roles for JNK and Aurora A in regulating the association of WDR62 with spindle microtubules. *J Cell Sci, 128*(3), 527-540.

Liu, H., Sadygov, R. G., & Yates, J. R., 3rd. (2004a). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem, 76*(14), 4193-4201.

Liu, J., & McLuckey, S. A. (2012). Electron Transfer Dissociation: Effects of Cation Charge State on Product Partitioning in Ion/Ion Electron Transfer to Multiply Protonated Polypeptides. *Int J Mass Spectrom, 330-332*, 174-181.

Liu, Q., Kaneko, S., Yang, L., Feldman, R. I., Nicosia, S. V., Chen, J., & Cheng, J. Q. (2004b). Aurora-A abrogation of p53 DNA binding and transactivation activity by phosphorylation of serine 215. *J Biol Chem, 279*(50), 52175-52182.

Liu, X. S., Song, B., Elzey, B. D., Ratliff, T. L., Konieczny, S. F., Cheng, L., Ahmad, N., & Liu, X. (2011). Polo-like kinase 1 facilitates loss of Pten tumor suppressor-induced prostate cancer formation. *J Biol Chem, 286*(41), 35795-35800.

Liu, Y., Gupta, G. D., Barnabas, D. D., Agircan, F. G., Mehmood, S., Wu, D., Coyaud, E., Johnson, C. M., McLaughlin, S. H., Andreeva, A., Freund, S. M. V., Robinson, C. V., Cheung, S. W. T., Raught, B., Pelletier, L., & van Breugel, M. (2018). Direct binding of CEP85 to STIL ensures robust PLK4 activation and efficient centriole assembly. *Nat Commun, 9*(1), 1731.

Lopes, C. A., Jana, S. C., Cunha-Ferreira, I., Zitouni, S., Bento, I., Duarte, P., Gilberto, S., Freixo, F., Guerrero, A., Francia, M., Lince-Faria, M., Carneiro, J., & Bettencourt-Dias, M. (2015). PLK4 trans-Autoactivation Controls Centriole Biogenesis in Space. *Dev Cell, 35*(2), 222-235.

Losada, A., Hirano, M., & Hirano, T. (2002). Cohesin release is required for sister chromatid resolution, but not for condensin-mediated compaction, at the onset of mitosis. *Genes Dev, 16*(23), 3004-3016.

Lowery, D. M., Lim, D., & Yaffe, M. B. (2005). Structure and function of Polo-like kinases. *Oncogene, 24*(2), 248-259.

Lui, C., Mok, M. T., & Henderson, B. R. (2016). Characterization of Adenomatous Polyposis Coli Protein Dynamics and Localization at the Centrosome. *Cancers (Basel), 8*(5).

Lundgren, D. H., Hwang, S. I., Wu, L., & Han, D. K. (2010). Role of spectral counting in quantitative proteomics. *Expert Rev Proteomics, 7*(1), 39-53.

Lunn, C. L., Chrivia, J. C., & Baldassare, J. J. (2010). Activation of Cdk2/Cyclin E complexes is dependent on the origin of replication licensing factor Cdc6 in mammalian cells. *Cell Cycle, 9*(22), 4533-4541.

Macmillan, J. C., Hudson, J. W., Bull, S., Dennis, J. W., & Swallow, C. J. (2001). Comparative expression of the mitotic regulators SAK and PLK in colorectal cancer. *Ann Surg Oncol, 8*(9), 729-740.

Macurek, L., Lindqvist, A., Lim, D., Lampson, M. A., Klompmaker, R., Freire, R., Clouin, C., Taylor, S. S., Yaffe, M. B., & Medema, R. H. (2008). Polo-like

kinase-1 is activated by aurora A to promote checkpoint recovery. *Nature, 455*(7209), 119-123.

Mailand, N., & Diffley, J. F. (2005). CDKs promote DNA replication origin licensing in human cells by protecting Cdc6 from APC/C-dependent proteolysis. *Cell, 122*(6), 915-926.

Makarov, A. (2000). Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem, 72*(6), 1156-1162.

Malumbres, M., & Barbacid, M. (2005). Mammalian cyclin-dependent kinases. *Trends Biochem Sci, 30*(11), 630-641.

Manfredi, M. G., Ecsedy, J. A., Chakravarty, A., Silverman, L., Zhang, M., Hoar, K. M., Stroud, S. G., Chen, W., Shinde, V., Huck, J. J., Wysong, D. R., Janowick, D. A., Hyer, M. L., Leroy, P. J., Gershman, R. E., Silva, M. D., Germanos, M. S., Bolen, J. B., Claiborne, C. F., & Sells, T. B. (2011). Characterization of Alisertib (MLN8237), an investigational small-molecule inhibitor of aurora A kinase using novel in vivo pharmacodynamic assays. *Clin Cancer Res, 17*(24), 7614-7624.

Mann, M., Hendrickson, R. C., & Pandey, A. (2001). Analysis of proteins and proteomes by mass spectrometry. *Annu Rev Biochem, 70*, 437-473.

Manning, G., Whyte, D. B., Martinez, R., Hunter, T., & Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science, 298*(5600), 1912-1934.

Manning, J. A., Shalini, S., Risk, J. M., Day, C. L., & Kumar, S. (2010). A direct interaction with NEDD1 regulates gamma-tubulin recruitment to the centrosome. *PLoS One, 5*(3), e9618.

Marcantonio, M., Trost, M., Courcelles, M., Desjardins, M., & Thibault, P. (2008). Combined enzymatic and data mining approaches for comprehensive phosphoproteome analyses: application to cell signaling events of interferon-gamma-stimulated macrophages. *Mol Cell Proteomics, 7*(4), 645-660.

March, R. E. (2012). Quadrupole Ion Trap Mass Spectrometer Update based on the original article by Raymond March, Encyclopedia of Analytical Chemistry, © 2000, John Wiley & Sons, Ltd. *Encyclopedia of Analytical Chemistry*.

Mardin, B. R., Agircan, F. G., Lange, C., & Schiebel, E. (2011). Plk1 controls the Nek2A-PP1gamma antagonism in centrosome disjunction. *Curr Biol, 21*(13), 1145-1151.

Martin, C. A., Ahmad, I., Klingseisen, A., Hussain, M. S., Bicknell, L. S., Leitch, A., Nurnberg, G., Toliat, M. R., Murray, J. E., Hunt, D., Khan, F., Ali, Z., Tinschert, S., Ding, J., Keith, C., Harley, M. E., Heyn, P., Muller, R., Hoffmann, I., Cormier-Daire, V., Dollfus, H., Dupuis, L., Bashamboo, A., McElreavey, K., Kariminejad, A., Mendoza-Londono, R., Moore, A. T., Saggar, A., Schlechter, C., Weleber, R., Thiele, H., Altmuller, J., Hohne, W., Hurles, M. E., Noegel, A. A., Baig, S. M., Nurnberg, P., & Jackson, A. P. (2014). Mutations in PLK4, encoding a master regulator of centriole biogenesis, cause microcephaly, growth failure and retinopathy. *Nat Genet, 46*(12), 1283-1292.

Marumoto, T., Hirota, T., Morisaki, T., Kunitoku, N., Zhang, D., Ichikawa, Y., Sasayama, T., Kuninaka, S., Mimori, T., Tamaki, N., Kimura, M., Okano, Y., &

Saya, H. (2002). Roles of aurora-A kinase in mitotic entry and G2 checkpoint in mammalian cells. *Genes Cells, 7*(11), 1173-1182.

Marumoto, T., Honda, S., Hara, T., Nitta, M., Hirota, T., Kohmura, E., & Saya, H. (2003). Aurora-A kinase maintains the fidelity of early and late mitotic events in HeLa cells. *J Biol Chem, 278*(51), 51786-51795.

Marx, H., Lemeer, S., Schliep, J. E., Matheron, L., Mohammed, S., Cox, J., Mann, M., Heck, A. J., & Kuster, B. (2013). A large synthetic peptide and phosphopeptide reference library for mass spectrometry-based proteomics. *Nat Biotechnol, 31*(6), 557-564.

Mason, J. M., Lin, D. C., Wei, X., Che, Y., Yao, Y., Kiarash, R., Cescon, D. W., Fletcher, G. C., Awrey, D. E., Bray, M. R., Pan, G., & Mak, T. W. (2014). Functional characterization of CFI-400945, a Polo-like kinase 4 inhibitor, as a potential anticancer agent. *Cancer Cell, 26*(2), 163-176.

Masters, S. C. (2004). Co-immunoprecipitation from transfected cells. *Methods Mol Biol, 261*, 337-350.

McAlister, G. C., Russell, J. D., Rumachik, N. G., Hebert, A. S., Syka, J. E., Geer, L. Y., Westphall, M. S., Pagliarini, D. J., & Coon, J. J. (2012). Analysis of the acidic proteome with negative electron-transfer dissociation mass spectrometry. *Anal Chem, 84*(6), 2875-2882.

McNulty, D. E., & Annan, R. S. (2008). Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. *Mol Cell Proteomics, 7*(5), 971-980.

Meraldi, P., Honda, R., & Nigg, E. A. (2002). Aurora-A overexpression reveals tetraploidization as a major route to centrosome amplification in p53-/- cells. *Embo j, 21*(4), 483-492.

Mori, Y., Inoue, Y., Tanaka, S., Doda, S., Yamanaka, S., Fukuchi, H., & Terada, Y. (2015a). Cep169, a Novel Microtubule Plus-End-Tracking Centrosomal Protein, Binds to CDK5RAP2 and Regulates Microtubule Stability. *PLoS One, 10*(10), e0140968.

Mori, Y., Inoue, Y., Taniyama, Y., Tanaka, S., & Terada, Y. (2015b). Phosphorylation of the centrosomal protein, Cep169, by Cdk1 promotes its dissociation from centrosomes in mitosis. *Biochem Biophys Res Commun, 468*(4), 642-646.

Moyer, T. C., Clutario, K. M., Lambrus, B. G., Daggubati, V., & Holland, A. J. (2015). Binding of STIL to Plk4 activates kinase activity to promote centriole assembly. *J Cell Biol, 209*(6), 863-878.

Musa, M. (2013). *Cell Proliferation Study of Human Osteosarcoma Cell Line (U2OS) using Alamar Blue Assayand Live Cell Imaging* (Vol. 8).

Negroni, L., Claverol, S., Rosenbaum, J., Chevet, E., Bonneu, M., & Schmitter, J. M. (2012). Comparison of IMAC and MOAC for phosphopeptide enrichment by column chromatography. *J Chromatogr B Analyt Technol Biomed Life Sci, 891-892*, 109-112.

Neville, D. C., Rozanas, C. R., Price, E. M., Gruis, D. B., Verkman, A. S., & Townsend, R. R. (1997). Evidence for phosphorylation of serine 753 in CFTR using a novel metal-ion affinity resin and matrix-assisted laser desorption mass spectrometry. *Protein Sci, 6*(11), 2436-2445.

Niforou, K. M., Anagnostopoulos, A. K., Vougas, K., Kittas, C., Gorgoulis, V. G., & Tsangaris, G. T. (2008). The proteome profile of the human osteosarcoma U2OS cell line. *Cancer Genomics Proteomics, 5*(1), 63-78.

Nigg, E. A. (2001). Mitotic kinases as regulators of cell division and its checkpoints. *Nat Rev Mol Cell Biol, 2*(1), 21-32.

Nishi, H., Shaytan, A., & Panchenko, A. R. (2014). Physicochemical mechanisms of protein regulation by phosphorylation. *Front Genet, 5*, 270.

Nishimura, M., Shin, M. S., Singhirunnusorn, P., Suzuki, S., Kawanishi, M., Koizumi, K., Saiki, I., & Sakurai, H. (2009). TAK1-mediated serine/threonine phosphorylation of epidermal growth factor receptor via p38/extracellular signal-regulated kinase: NF-{kappa}B-independent survival pathways in tumor necrosis factor alpha signaling. *Mol Cell Biol, 29*(20), 5529-5539.

Nowakowski, J., Cronin, C. N., McRee, D. E., Knuth, M. W., Nelson, C. G., Pavletich, N. P., Rogers, J., Sang, B. C., Scheibe, D. N., Swanson, R. V., & Thompson, D. A. (2002). Structures of the cancer-related Aurora-A, FAK, and EphA2 protein kinases from nanovolume crystallography. *Structure, 10*(12), 1659-1667.

O'Hare, T., Shakespeare, W. C., Zhu, X., Eide, C. A., Rivera, V. M., Wang, F., Adrian, L. T., Zhou, T., Huang, W. S., Xu, Q., Metcalf, C. A., 3rd, Tyner, J. W., Loriaux, M. M., Corbin, A. S., Wardwell, S., Ning, Y., Keats, J. A., Wang, Y., Sundaramoorthi, R., Thomas, M., Zhou, D., Snodgrass, J., Commodore, L., Sawyer, T. K., Dalgarno, D. C., Deininger, M. W., Druker, B. J., & Clackson, T. (2009). AP24534, a pan-BCR-ABL inhibitor for chronic myeloid leukemia, potently inhibits the T315I mutant and overcomes mutation-based resistance. *Cancer Cell, 16*(5), 401-412.

Ohta, M., Ashikawa, T., Nozaki, Y., Kozuka-Hata, H., Goto, H., Inagaki, M., Oyama, M., & Kitagawa, D. (2014). Direct interaction of Plk4 with STIL ensures formation of a single procentriole per parental centriole. *Nat Commun, 5*, 5267.

Okuda, M., Horn, H. F., Tarapore, P., Tokuyama, Y., Smulian, A. G., Chan, P. K., Knudsen, E. S., Hofmann, I. A., Snyder, J. D., Bove, K. E., & Fukasawa, K. (2000). Nucleophosmin/B23 is a target of CDK2/cyclin E in centrosome duplication. *Cell, 103*(1), 127-140.

Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., & Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell, 127*(3), 635-648.

Olsen, J. V., Macek, B., Lange, O., Makarov, A., Horning, S., & Mann, M. (2007). Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods, 4*(9), 709-712.

Olsen, J. V., & Mann, M. (2013). Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol Cell Proteomics, 12*(12), 3444-3452.

Ong, S.-E., & Mann, M. (2007). A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nature Protocols, 1*, 2650.

Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., & Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics, 1*(5), 376-386.

Ong, S. E., Kratchmarova, I., & Mann, M. (2003). Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J Proteome Res, 2*(2), 173-181.

Ou, B., Zhao, J., Guan, S., Wangpu, X., Zhu, C., Zong, Y., Ma, J., Sun, J., Zheng, M., Feng, H., & Lu, A. (2016). Plk2 promotes tumor growth and inhibits apoptosis by targeting Fbxw7/Cyclin E in colorectal cancer. *Cancer Lett, 380*(2), 457-466.

Palumbo, A. M., Tepe, J. J., & Reid, G. E. (2008). Mechanistic insights into the multistage gas-phase fragmentation behavior of phosphoserine- and phosphothreonine-containing peptides. *J Proteome Res, 7*(2), 771-779.

Pao, K. C., Wood, N. T., Knebel, A., Rafie, K., Stanley, M., Mabbitt, P. D., Sundaramoorthy, R., Hofmann, K., van Aalten, D. M. F., & Virdee, S. (2018). Activity-based E3 ligase profiling uncovers an E3 ligase with esterification activity. *Nature, 556*(7701), 381-385.

Perkins, D. N., Pappin, D. J., Creasy, D. M., & Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis, 20*(18), 3551-3567.

Petretti, C., Savoian, M., Montembault, E., Glover, D. M., Prigent, C., & Giet, R. (2006). The PITSLRE/CDK11p58 protein kinase promotes centrosome maturation and bipolar spindle formation. *EMBO Rep, 7*(4), 418-424.

Pitteri, S. J., Chrisman, P. A., Hogan, J. M., & McLuckey, S. A. (2005). Electron transfer ion/ion reactions in a three-dimensional quadrupole ion trap: reactions of doubly and triply protonated peptides with SO2*. *Anal Chem, 77*(6), 1831-1839.

Plotnikova, O. V., Nikonova, A. S., Loskutov, Y. V., Kozyulina, P. Y., Pugacheva, E. N., & Golemis, E. A. (2012). Calmodulin activation of Aurora-A kinase (AURKA) is required during ciliary disassembly and in mitosis. *Mol Biol Cell, 23*(14), 2658-2670.

Plotnikova, O. V., Pugacheva, E. N., Dunbrack, R. L., & Golemis, E. A. (2010). Rapid calcium-dependent activation of Aurora-A kinase. *Nat Commun, 1*, 64.

Poliakova, M., Aebersold, D. M., Zimmer, Y., & Medova, M. (2018). The relevance of tyrosine kinase inhibitors for global metabolic pathways in cancer. *Mol Cancer, 17*(1), 27.

Poss, Z. C., Ebmeier, C. C., Odell, A. T., Tangpeerachaikul, A., Lee, T., Pelish, H. E., Shair, M. D., Dowell, R. D., Old, W. M., & Taatjes, D. J. (2016). Identification of Mediator Kinase Substrates in Human Cells using Cortistatin A and Quantitative Phosphoproteomics. *Cell Rep, 15*(2), 436-450.

Puklowski, A., Homsi, Y., Keller, D., May, M., Chauhan, S., Kossatz, U., Grunwald, V., Kubicka, S., Pich, A., Manns, M. P., Hoffmann, I., Gonczy, P., & Malek, N. P. (2011). The SCF-FBXW5 E3-ubiquitin ligase is regulated by PLK4 and targets HsSAS-6 to control centrosome duplication. *Nat Cell Biol, 13*(8), 1004-1009.

Qin, W. W., Sang, C. Y., Zhang, L. L., Wei, W., Tian, H. Z., Liu, H. X., Chen, S. W., & Hui, L. (2015). Synthesis and biological evaluation of 2,4-diaminopyrimidines as selective Aurora A kinase inhibitors. *Eur J Med Chem, 95*, 174-184.

Ren, B., Cam, H., Takahashi, Y., Volkert, T., Terragni, J., Young, R. A., & Dynlacht, B. D. (2002). E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes Dev, 16*(2), 245-256.

Rieder, C. L., Davison, E. A., Jensen, L. C., Cassimeris, L., & Salmon, E. D. (1986). Oscillatory movements of monooriented chromosomes and their position relative to the spindle pole result from the ejection properties of the aster and half-spindle. *J Cell Biol, 103*(2), 581-591.

Riley, N. M., & Coon, J. J. (2018). The Role of Electron Transfer Dissociation in Modern Proteomics. *Anal Chem, 90*(1), 40-64.

Riley, N. M., Mullen, C., Weisbrod, C. R., Sharma, S., Senko, M. W., Zabrouskov, V., Westphall, M. S., Syka, J. E., & Coon, J. J. (2016). Enhanced Dissociation of Intact Proteins with High Capacity Electron Transfer Dissociation. *J Am Soc Mass Spectrom, 27*(3), 520-531.

Riley, N. M., Rush, M. J., Rose, C. M., Richards, A. L., Kwiecien, N. W., Bailey, D. J., Hebert, A. S., Westphall, M. S., & Coon, J. J. (2015). The Negative Mode Proteome with Activated Ion Negative Electron Transfer Dissociation (AI-NETD). *Mol Cell Proteomics, 14*(10), 2644-2660.

Rivers, J., Simpson, D. M., Robertson, D. H., Gaskell, S. J., & Beynon, R. J. (2007). Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. *Mol Cell Proteomics, 6*(8), 1416-1427.

Roepstorff, P., & Fohlman, J. (1984). Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom, 11*(11), 601.

Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., & Pappin, D. J. (2004). Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics, 3*(12), 1154-1169.

Rubin, S. M. (2013). Deciphering the retinoblastoma protein phosphorylation code. *Trends Biochem Sci, 38*(1), 12-19.

Rubin, S. M., Gall, A. L., Zheng, N., & Pavletich, N. P. (2005). Structure of the Rb C-terminal domain bound to E2F1-DP1: a mechanism for phosphorylation-induced E2F release. *Cell, 123*(6), 1093-1106.

Rudolf, A. F., Skovgaard, T., Knapp, S., Jensen, L. J., & Berthelsen, J. (2014). A comparison of protein kinases inhibitor screening methods using both enzymatic activity and binding affinity determination. *PLoS One, 9*(6), e98800.

Rudolph, D., Steegmaier, M., Hoffmann, M., Grauert, M., Baum, A., Quant, J., Haslinger, C., Garin-Chesa, P., & Adolf, G. R. (2009). BI 6727, a Polo-like kinase inhibitor with improved pharmacokinetic profile and broad antitumor activity. *Clin Cancer Res, 15*(9), 3094-3102.

S. Wilm, M., & Mann, M. (1994). *Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last?* (Vol. 136).

Salina, D., Bodoor, K., Eckley, D. M., Schroer, T. A., Rattner, J. B., & Burke, B. (2002). Cytoplasmic dynein as a facilitator of nuclear envelope breakdown. *Cell, 108*(1), 97-107.

Santamaria, A., Wang, B., Elowe, S., Malik, R., Zhang, F., Bauer, M., Schmidt, A., Sillje, H. H., Korner, R., & Nigg, E. A. (2011). The Plk1-dependent phosphoproteome of the early mitotic spindle. *Mol Cell Proteomics, 10*(1), M110.004457.

Sasai, K., Katayama, H., Hawke, D. H., & Sen, S. (2016). Aurora-C Interactions with Survivin and INCENP Reveal Shared and Distinct Features Compared with Aurora-B Chromosome Passenger Protein Complex. *PLoS One, 11*(6), e0157305.

Sato, K., Shin, M. S., Sakimura, A., Zhou, Y., Tanaka, T., Kawanishi, M., Kawasaki, Y., Yokoyama, S., Koizumi, K., Saiki, I., & Sakurai, H. (2013). Inverse correlation between Thr-669 and constitutive tyrosine phosphorylation in the asymmetric epidermal growth factor receptor dimer conformation. *Cancer Sci, 104*(10), 1315-1322.

Savitski, M. M., Lemeer, S., Boesche, M., Lang, M., Mathieson, T., Bantscheff, M., & Kuster, B. (2011). Confident phosphorylation site localization using the Mascot Delta Score. *Mol Cell Proteomics, 10*(2), M110.003830.

Schafer, K. A. (1998). The cell cycle: a review. *Vet Pathol, 35*(6), 461-478.

Schenk, P. W., & Snaar-Jagalska, B. E. (1999). Signal perception and transduction: the role of protein kinases. *Biochim Biophys Acta, 1449*(1), 1-24.

Schlosser, A., & Lehmann, W. D. (2000). Five-membered ring formation in unimolecular reactions of peptides: a key structural element controlling low-energy collision-induced dissociation of peptides. *J Mass Spectrom, 35*(12), 1382-1390.

Schroeder, M. J., Shabanowitz, J., Schwartz, J. C., Hunt, D. F., & Coon, J. J. (2004). A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. *Anal Chem, 76*(13), 3590-3598.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., & Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature, 473*(7347), 337-342.

Schwartz, D., & Gygi, S. P. (2005). An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat Biotechnol, 23*(11), 1391-1398.

Schwartz, J. C., Senko, M. W., & Syka, J. E. (2002). A two-dimensional quadrupole ion trap mass spectrometer. *J Am Soc Mass Spectrom, 13*(6), 659-669.

Scott, I. S., Morris, L. S., Bird, K., Davies, R. J., Vowler, S. L., Rushbrook, S. M., Marshall, A. E., Laskey, R. A., Miller, R., Arends, M. J., & Coleman, N. (2003). A novel immunohistochemical method to estimate cell-cycle phase distribution in archival tissue: implications for the prediction of outcome in colorectal cancer. *J Pathol, 201*(2), 187-197.

Scutt, P. J., Chu, M. L., Sloane, D. A., Cherry, M., Bignell, C. R., Williams, D. H., & Eyers, P. A. (2009). Discovery and exploitation of inhibitor-resistant aurora and polo kinase mutants for the analysis of mitotic networks. *J Biol Chem, 284*(23), 15880-15893.

Sebastian, B., Kakizuka, A., & Hunter, T. (1993). Cdc25M2 activation of cyclin-dependent kinases by dephosphorylation of threonine-14 and tyrosine-15. *Proc Natl Acad Sci U S A, 90*(8), 3521-3524.

Seki, A., Coppinger, J. A., Jang, C. Y., Yates, J. R., & Fang, G. (2008). Bora and the kinase Aurora a cooperatively activate the kinase Plk1 and control mitotic entry. *Science, 320*(5883), 1655-1658.

Senko, M. W., Remes, P.M., Canterbury, J. D., Mathur, R., Song, Q., Eiluk, S. M., Mullen, C., Earley, L.,Hardman, M., Blethrow, J. D., Bui, H., Specht, A., Lange, O., Denisov, E., Makarov, A., Horning, S. & Zabrouskov, V. (2013). Novel Parallelized Quadrupole/Linear Ion Trap/Orbitrap Tribrid Mass Spectrometer Improving Proteome Coverage and Peptide Identification Rates. *Anal Chem*, *85*(24), 11710-11714.

Shah, N. P., Tran, C., Lee, F. Y., Chen, P., Norris, D., & Sawyers, C. L. (2004). Overriding imatinib resistance with a novel ABL kinase inhibitor. *Science, 305*(5682), 399-401.

Shinmura, K., Kurabe, N., Goto, M., Yamada, H., Natsume, H., Konno, H., & Sugimura, H. (2014). PLK4 overexpression and its effect on centrosome regulation and chromosome stability in human gastric cancer. *Mol Biol Rep, 41*(10), 6635-6644.

Shrestha, A., Hamilton, G., O'Neill, E., Knapp, S., & Elkins, J. M. (2012). Analysis of conditions affecting auto-phosphorylation of human kinases during expression in bacteria. *Protein Expr Purif, 81*(1), 136-143.

Shu, X., Fry, A. M., Tulloch, B., Manson, F. D., Crabb, J. W., Khanna, H., Faragher, A. J., Lennon, A., He, S., Trojan, P., Giessl, A., Wolfrum, U., Vervoort, R., Swaroop, A., & Wright, A. F. (2005). RPGR ORF15 isoform co-localizes with RPGRIP1 at centrioles and basal bodies and interacts with nucleophosmin. *Hum Mol Genet, 14*(9), 1183-1197.

Sillibourne, J. E., Delaval, B., Redick, S., Sinha, M., & Doxsey, S. J. (2007). Chromatin remodeling proteins interact with pericentrin to regulate centrosome integrity. *Mol Biol Cell, 18*(9), 3667-3680.

Sillibourne, J. E., Tack, F., Vloemans, N., Boeckx, A., Thambirajah, S., Bonnet, P., Ramaekers, F. C., Bornens, M., & Grand-Perret, T. (2010). Autophosphorylation of polo-like kinase 4 and its role in centriole duplication. *Mol Biol Cell, 21*(4), 547-561.

Singh, S. V., Herman-Antosiewicz, A., Singh, A. V., Lew, K. L., Srivastava, S. K., Kamath, R., Brown, K. D., Zhang, L., & Baskaran, R. (2004). Sulforaphane-induced G2/M phase cell cycle arrest involves checkpoint kinase 2-mediated phosphorylation of cell division cycle 25C. *J Biol Chem, 279*(24), 25813-25822.

Slevin, L. K., Nye, J., Pinkerton, D. C., Buster, D. W., Rogers, G. C., & Slep, K. C. (2012). The structure of the plk4 cryptic polo box reveals two tandem polo boxes required for centriole duplication. *Structure, 20*(11), 1905-1917.

Sloane, D. A., Trikic, M. Z., Chu, M. L., Lamers, M. B., Mason, C. S., Mueller, I., Savory, W. J., Williams, D. H., & Eyers, P. A. (2010). Drug-resistant aurora A mutants for cellular target validation of the small molecule kinase inhibitors MLN8054 and MLN8237. *ACS Chem Biol, 5*(6), 563-576.

Smith, M. R., Wilson, M. L., Hamanaka, R., Chase, D., Kung, H., Longo, D. L., & Ferris, D. K. (1997). Malignant transformation of mammalian cells initiated by constitutive expression of the polo-like kinase. *Biochem Biophys Res Commun, 234*(2), 397-405.

Snyder, D. T., Peng, W.-P., & Cooks, R. G. (2017). Resonance methods in quadrupole ion traps. *Chemical Physics Letters, 668*, 69-89.

Solari, F. A., Dell'Aica, M., Sickmann, A., & Zahedi, R. P. (2015). Why phosphoproteomics is still a challenge. *Mol Biosyst, 11*(6), 1487-1493.

Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M. F., Piwnica-Worms, H., & Cantley, L. C. (1994). Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr Biol, 4*(11), 973-982.

Sonnen, K. F., Gabryjonczyk, A. M., Anselm, E., Stierhof, Y. D., & Nigg, E. A. (2013). Human Cep192 and Cep152 cooperate in Plk4 recruitment and centriole duplication. *J Cell Sci, 126*(Pt 14), 3223-3233.

Spitaler, M., Villunger, A., Grunicke, H., & Uberall, F. (2000). Unique structural and functional properties of the ATP-binding domain of atypical protein kinase C-iota. *J Biol Chem, 275*(43), 33289-33296.

Stafford, G., Kelley, P. E., Syka, J. E. P., Reynolds, W. E., & Todd, J. (1984). *Recent Improvements in and Analytical Applications of Advanced Ion Trap Technology* (Vol. 60).

Steen, H., Jebanathirajah, J. A., Rush, J., Morrice, N., & Kirschner, M. W. (2006). Phosphorylation analysis by mass spectrometry: myths, facts, and the consequences for qualitative and quantitative measurements. *Mol Cell Proteomics, 5*(1), 172-181.

Summerfield, S. G., & Gaskell, S. J. (1997). Fragmentation efficiencies of peptide ions following low energy collisional activation. *International Journal of Mass Spectrometry and Ion Processes, 165-166*, 509-521.

Sunkel, C. E., & Glover, D. M. (1988). polo, a mitotic mutant of Drosophila displaying abnormal spindle poles. *J Cell Sci, 89 (Pt 1)*, 25-38.

Swaffer, M. P., Jones, A. W., Flynn, H. R., Snijders, A. P., & Nurse, P. (2016). CDK Substrate Phosphorylation and Ordering the Cell Cycle. *Cell, 167*(7), 1750-1761.e1716.

Swallow, C. J., Ko, M. A., Siddiqui, N. U., Hudson, J. W., & Dennis, J. W. (2005). Sak/Plk4 and mitotic fidelity. *Oncogene, 24*(2), 306-312.

Swaney, D. L., McAlister, G. C., Wirtala, M., Schwartz, J. C., Syka, J. E., & Coon, J. J. (2007). Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors. *Anal Chem, 79*(2), 477-485.

Syed, N., Smith, P., Sullivan, A., Spender, L. C., Dyer, M., Karran, L., O'Nions, J., Allday, M., Hoffmann, I., Crawford, D., Griffin, B., Farrell, P. J., & Crook, T. (2006). Transcriptional silencing of Polo-like kinase 2 (SNK/PLK2) is a frequent event in B-cell malignancies. *Blood, 107*(1), 250-256.

Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., & Hunt, D. F. (2004). Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A, 101*(26), 9528-9533.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J., & von Mering, C. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res, 43*(Database issue), D447-452.

Tabb, D. L., Huang, Y., Wysocki, V. H., & Yates, J. R., 3rd. (2004). Influence of basic residue content on fragment ion peak intensities in low-energy collision-induced dissociation spectra of peptides. *Anal Chem, 76*(5), 1243-1248.

Taguchi, S., Honda, K., Sugiura, K., Yamaguchi, A., Furukawa, K., & Urano, T. (2002). Degradation of human Aurora-A protein kinase is mediated by hCdh1. *FEBS Lett, 519*(1-3), 59-65.

Taus, T., Kocher, T., Pichler, P., Paschke, C., Schmidt, A., Henrich, C., & Mechtler, K. (2011). Universal and confident phosphorylation site localization using phosphoRS. *J Proteome Res, 10*(12), 5354-5362.

Taylor, G. I., & McEwan, A. D. (1965). The stability of a horizontal fluid interface in a vertical electric field. *Journal of Fluid Mechanics, 22*(1), 1-15.

Tennant, D. A., Duran, R. V., & Gottlieb, E. (2010). Targeting metabolic transformation for cancer therapy. *Nat Rev Cancer, 10*(4), 267-277.

Thingholm, T. E., Jørgensen, T. J. D., Jensen, O. N., & Larsen, M. R. (2006). Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nature Protocols, 1*, 1929.

Tholey, A., Reed, J., & Lehmann, W. D. (1999). Electrospray tandem mass spectrometric studies of phosphopeptides and phosphopeptide analogues. *J Mass Spectrom, 34*(2), 117-123.

Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A. K., & Hamon, C. (2003). Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem, 75*(8), 1895-1904.

Thomson, B., & Iribarne, J. V. J. (1979). *Field Induced Ion Evaporation from Liquid Surfaces at Atmospheric Pressure* (Vol. 71).

Travers, T., Shao, H., Joughin, B. A., Lauffenburger, D. A., Wells, A., & Camacho, C. J. (2015). Tandem phosphorylation within an intrinsically disordered region regulates ACTN4 function. *Sci Signal, 8*(378), ra51.

Treiber, D. K., & Shah, N. P. (2013). Ins and outs of kinase DFG motifs. *Chem Biol, 20*(6), 745-746.

Trost, M., Bridon, G., Desjardins, M., & Thibault, P. (2010). Subcellular phosphoproteomics. *Mass Spectrom Rev, 29*(6), 962-990.

Tseng, T. C., Chen, S. H., Hsu, Y. P., & Tang, T. K. (1998). Protein kinase profile of sperm and eggs: cloning and characterization of two novel testis-specific protein kinases (AIE1, AIE2) related to yeast and fly chromosome segregation regulators. *DNA Cell Biol, 17*(10), 823-833.

Tsutsumi, M., Yokoi, S., Miya, F., Miyata, M., Kato, M., Okamoto, N., Tsunoda, T., Yamasaki, M., Kanemura, Y., Kosaki, K., Saitoh, S., & Kurahashi, H. (2016). Novel compound heterozygous variants in PLK4 identified in a patient with autosomal recessive microcephaly and chorioretinopathy. *Eur J Hum Genet, 24*(12), 1702-1706.

Tyler, R. K., Shpiro, N., Marquez, R., & Eyers, P. A. (2007). VX-680 inhibits Aurora A and Aurora B kinase activity in human cells. *Cell Cycle, 6*(22), 2846-2854.

V. Iribarne, J., & Thomson, B. (1976). *On the Evaporation of Small Ions from Charged Droplets* (Vol. 64).

Vedel, F., & André, J. (1984). Influence of space charge on the computed statistical properties of stored ions cooled by a buffer gas in a quadrupole rf trap. *Physical Review A, 29*(4), 2098-2101.

Waizenegger, I. C., Hauf, S., Meinke, A., & Peters, J. M. (2000). Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase. *Cell, 103*(3), 399-410.

Walter, A. O., Seghezzi, W., Korver, W., Sheung, J., & Lees, E. (2000). The mitotic serine/threonine kinase Aurora2/AIK is regulated by phosphorylation and degradation. *Oncogene, 19*(42), 4906-4916.

Wang, W., Budhu, A., Forgues, M., & Wang, X. W. (2005). Temporal and spatial control of nucleophosmin by the Ran-Crm1 complex in centrosome duplication. *Nat Cell Biol, 7*(8), 823-830.

Wang, G., Jiang, Q. & Zhang, C. (2014). The role of mitotic kinases in coupling the centrosome cycle with the assembly of the mitotic spindle. *J. Cell. Sci, 127*, 4111-4122.

Welburn, J. P., Vleugel, M., Liu, D., Yates, J. R., 3rd, Lampson, M. A., Fukagawa, T., & Cheeseman, I. M. (2010). Aurora B phosphorylates spatially distinct targets to differentially regulate the kinetochore-microtubule interface. *Mol Cell, 38*(3), 383-392.

Wiese, H., Kuhlmann, K., Wiese, S., Stoepel, N. S., Pawlas, M., Meyer, H. E., Stephan, C., Eisenacher, M., Drepper, F., & Warscheid, B. (2014). Comparison of alternative MS/MS and bioinformatics approaches for confident phosphorylation site localization. *J Proteome Res, 13*(2), 1128-1137.

Wilhelm, M., Schlegl, J., Hahne, H., Gholami, A. M., Lieberenz, M., Savitski, M. M., Ziegler, E., Butzmann, L., Gessulat, S., Marx, H., Mathieson, T., Lemeer, S., Schnatbaum, K., Reimer, U., Wenschuh, H., Mollenhauer, M., Slotta-Huspenina, J., Boese, J. H., Bantscheff, M., Gerstmair, A., Faerber, F., & Kuster, B. (2014). Mass-spectrometry-based draft of the human proteome. *Nature, 509*(7502), 582-587.

Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., Golaz, O., Sanchez, J. C., Yan, J. X., Gooley, A. A., Hughes, G., Humphery-Smith, I., Williams, K. L., & Hochstrasser, D. F. (1996). From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y), 14*(1), 61-65.

Wilkinson, M. G., & Millar, J. B. (2000). Control of the eukaryotic cell cycle by MAP kinase signaling pathways. *Faseb j, 14*(14), 2147-2157.

Wilm, M. (2011). Principles of electrospray ionization. *Mol Cell Proteomics, 10*(7), M111.009407.

Wong, J., Lerrigo, R., Jang, C. Y., & Fang, G. (2008). Aurora A regulates the activity of HURP by controlling the accessibility of its microtubule-binding domain. *Mol Biol Cell, 19*(5), 2083-2091.

Wong, Y. L., Anzola, J. V., Davis, R. L., Yoon, M., Motamedi, A., Kroll, A., Seo, C. P., Hsia, J. E., Kim, S. K., Mitchell, J. W., Mitchell, B. J., Desai, A., Gahman, T. C., Shiau, A. K., & Oegema, K. (2015). Cell biology. Reversible centriole depletion with an inhibitor of Polo-like kinase 4. *Science, 348*(6239), 1155-1160.

Wu, J., Shakey, Q., Liu, W., Schuller, A., & Follettie, M. T. (2007). Global profiling of phosphopeptides by titania affinity enrichment. *J Proteome Res, 6*(12), 4684-4689.

Xie, S., Wu, H., Wang, Q., Cogswell, J. P., Husain, I., Conn, C., Stambrook, P., Jhanwar-Uniyal, M., & Dai, W. (2001). Plk3 functionally links DNA damage to

cell cycle arrest and apoptosis at least in part via the p53 pathway. *J Biol Chem, 276*(46), 43305-43312.

Xiong, Y., Zhang, H., & Beach, D. (1993). Subunit rearrangement of the cyclin-dependent kinases is associated with cellular transformation. *Genes Dev, 7*(8), 1572-1583.

Xu, D., Yao, Y., Lu, L., Costa, M., & Dai, W. (2010). Plk3 functions as an essential component of the hypoxia regulatory pathway by direct phosphorylation of HIF-1alpha. *J Biol Chem, 285*(50), 38944-38950.

Xu, X., Huang, S., Zhang, B., Huang, F., Chi, W., Fu, J., Wang, G., Li, S., Jiang, Q., & Zhang, C. (2017). DNA replication licensing factor Cdc6 and Plk4 kinase antagonistically regulate centrosome duplication via Sas-6. *Nat Commun, 8*, 15164.

Yang, F., Shen, Y., Camp, D. G., 2nd, & Smith, R. D. (2012). High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis. *Expert Rev Proteomics, 9*(2), 129-134.

Yang, J., Adamian, M., & Li, T. (2006). Rootletin interacts with C-Nap1 and may function as a physical linker between the pair of centrioles/basal bodies in cells. *Mol Biol Cell, 17*(2), 1033-1040.

Yang, Y., Bai, J., Shen, R., Brown, S. A., Komissarova, E., Huang, Y., Jiang, N., Alberts, G. F., Costa, M., Lu, L., Winkles, J. A., & Dai, W. (2008). Polo-like kinase 3 functions as a tumor suppressor and is a negative regulator of hypoxia-inducible factor-1 alpha under hypoxic conditions. *Cancer Res, 68*(11), 4077-4085.

Yim, H., Park, J. W., Woo, S. U., Kim, S. T., Liu, L., Lee, C. H., & Lee, S. K. (2013). Phosphorylation of Cdc6 at serine 74, but not at serine 106, drives translocation of Cdc6 to the cytoplasm. *J Cell Physiol, 228*(6), 1221-1228.

Zarkowska, T., & Mittnacht, S. (1997). Differential phosphorylation of the retinoblastoma protein by G1/S cyclin-dependent kinases. *J Biol Chem, 272*(19), 12738-12746.

Zekri, A., Lesan, V., Ghaffari, S. H., Tabrizi, M. H., & Modarressi, M. H. (2012). Gene amplification and overexpression of Aurora-C in breast and prostate cancer cell lines. *Oncol Res, 20*(5-6), 241-250.

Zhang, Y., Ficarro, S. B., Li, S., & Marto, J. A. (2009). Optimized Orbitrap HCD for quantitative analysis of phosphopeptides. *J Am Soc Mass Spectrom, 20*(8), 1425-1434.

Zhao, Z. S., Lim, J. P., Ng, Y. W., Lim, L., & Manser, E. (2005). The GIT-associated kinase PAK targets to the centrosome and regulates Aurora-A. *Mol Cell, 20*(2), 237-249.

Zhou, H., Di Palma, S., Preisinger, C., Peng, M., Polat, A. N., Heck, A. J., & Mohammed, S. (2013). Toward a comprehensive characterization of a human cancer cell phosphoproteome. *J Proteome Res, 12*(1), 260-271.

Zimmerman, W. C., & Erikson, R. L. (2007). Polo-like kinase 3 is required for entry into S phase. *Proc Natl Acad Sci U S A, 104*(6), 1847-1852.

Zitouni, S., Francia, M. E., Leal, F., Montenegro Gouveia, S., Nabais, C., Duarte, P., Gilberto, S., Brito, D., Moyer, T., Kandels-Lewis, S., Ohta, M., Kitagawa, D., Holland, A. J., Karsenti, E., Lorca, T., Lince-Faria, M., & Bettencourt-Dias, M.

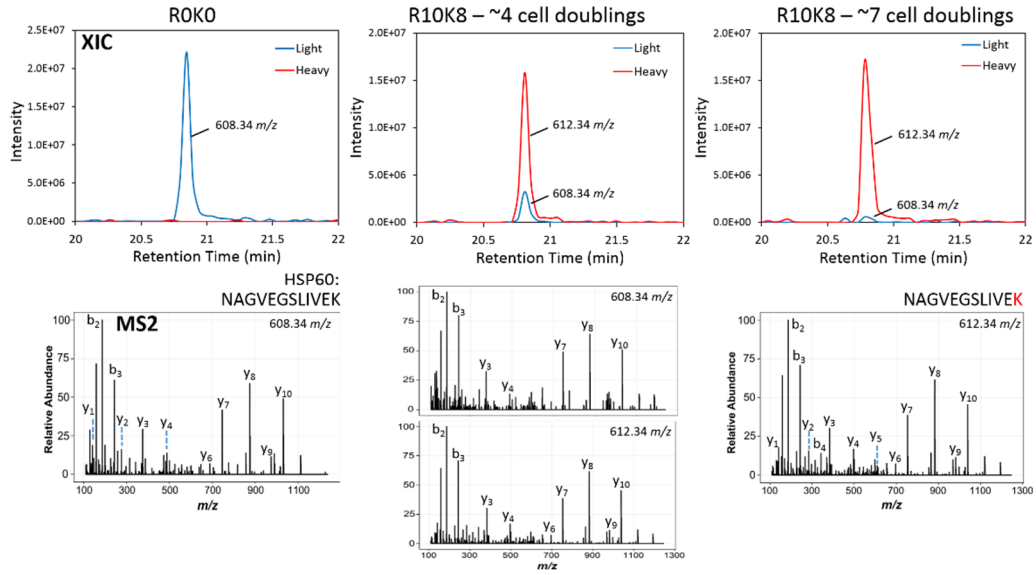
(2016). CDK1 Prevents Unscheduled PLK4-STIL Complex Assembly in Centriole Biogenesis. *Curr Biol, 26*(9), 1127-1137.

Zorba, A., Buosi, V., Kutter, S., Kern, N., Pontiggia, F., Cho, Y. J., & Kern, D. (2014). Molecular mechanism of Aurora A kinase autophosphorylation and its allosteric activation by TPX2. *Elife, 3*, e02667.

Zubarev, R. A., & Makarov, A. (2013). Orbitrap Mass Spectrometry. *Anal Chem, 85*(11), 5288-5296.

9. Appendix

Appendix 1



Heavy lysine ($^{13}C_6$ $^{15}N_2$) incorporation in U2OS cells. The extracted ion chromatograms (top panels) show the ion signals for the indicated peptide. The top left panel shows a peptide at *m/z* 608.34 and represents the presence of the peptide in its light form due to growth in 'light' media. MS/MS (bottom left) reveals that the peptide is from HSP60. Following 4 cell doublings in SILAC (R10K8) media, the same peptide is shown as a SILAC pair, with a mass difference of 8 Da. The 'heavy' peptide is at a higher intensity but still contains sufficient 'light' peptide to be identified by MS/MS (bottom, middle panels), reflecting partial incorporation of the label. At 7 cell doublings, no 'light' peptide can be identified reflecting full incorporation of the heavy label.

Appendix 2

List of centrosomal-associated phosphoproteins. Analysis of significantly downregulated (A) and upregulated (B) (adj. *p* value ≤0.075) phosphosites in DAVID revealed an enrichment of phosphoproteins localised at the centrosome.

A

Gene Name	pSite	PTM-score	ptmRS	Fold Change (Log2)	Adj. *p* value
NCKAP5L	Ser567	0.995	98.9	-1.119	0.069
CHD4	Ser531	0.976	NA	-0.931	0.074
PPP1R12A	Ser445	0.999	100	-0.834	0.050
NUMA1	Ser1991	1.000	100	-0.688	0.052
LZTS2	Ser311	1.000	100	-0.626	0.052
RANBP1	Ser60	1.000	100	-0.608	0.072
TSPYL2	Ser16	0.949	100	-0.545	0.062
ARHGEF10	Tyr239	1.000	100	-0.539	0.047
XRCC4	Ser256	1.000	100	-0.529	0.068
DHX9	Ser87	1.000	98.5	-0.416	0.047
MPLKIP	Ser40	0.999	NA	-0.384	0.067
APC	Ser1436	0.832	50.0	-0.282	0.075
CLIP1	Ser147	0.976	50.0	-0.275	0.068
WDR62	Ser33	0.901	99.7	-0.254	0.069

B

Gene Name	pSite	PTM-score	ptmRS	Fold Change (Log2)	Adj. *p* value
NPM1	Ser10	1.000	100	0.309	0.058
1	Ser280	0.999	100	0.334	0.072
RPGR	Ser961	1.000	100	0.397	0.067
DTL	Ser485	0.975	100	0.411	0.064
LEO1	Ser10	1.000	NA	0.415	0.064
PLK4	Ser421	0.999	100	0.878	0.044
PLK4	Ser665	0.968	100	1.224	0.044
PLK4	Ser817	1.000	100	1.876	0.047

Appendix 3

The following excel files are available upon request:

Table 9.1. Preparation of synthetic phosphopeptide library pools. Lists of the phosphopeptides included in each of the five pools created are shown. Phosphoisomers were separated to ensure the exact site of localisation could be identified by LC-MS/MS and downstream data processing.

Table 9.2. Bayesian statistical analysis of protein expression data in the WT PLK4 dataset.

Table 9.3. Bayesian statistical analysis of protein expression data in the G95R PLK4 dataset

Identified proteins were filtered in Perseus to remove 'contaminants', 'reverse hits' and those identified 'only by site'. In addition, proteins were filtered to include those identified in ≥3 bioreplicates. Bayesian statistical analysis was then performed using the LIMMA package in R, with the Benjamini-Hochberg multiple corrections procedure to generate adj. *p* values.

Table 9.4. Bayesian statistical analysis of phosphosite data in the WT PLK4 dataset.

Table 9.5. Bayesian statistical analysis of phosphosite data in the G95R PLK4 dataset.

Identified proteins were filtered in Perseus to remove 'contaminants', 'reverse hits' and those identified 'only by site'. The site table was expanded to separate phosphosites. In addition, proteins were filtered to include those identified in ≥3 bioreplicates and with a phosphosite localisation score ≥ 0.75. Bayesian statistical analysis was then performed using the LIMMA package in R, with the Benjamini-Hochberg multiple corrections procedure to generate adj. *p* values.

Table 9.6. DAVID enrichment analysis of significantly regulated proteins in the WT PLK4 dataset.

Table 9.7. DAVID enrichment analysis of significantly regulated proteins in the G95R PLK4 dataset.

Table 9.8. DAVID enrichment analysis of significantly regulated phosphosites in the WT PLK4 dataset.

Table 9.9. DAVID enrichment analysis of significantly regulated phosphosites in the G95R PLK4 dataset.

Analysis of significantly regulated proteins and phosphosites (adj. *p* value ≤0.075 for WT and adj. *p* value ≤0.075 for G95R) in DAVID revealed enrichment at specific cellular compartments, in biological processes, molecular functions and KEGG pathways.

Appendix 4

Publication freely available online:

Ferries, S., Perkins, S., Brownridge, P. J., Campbell, A., Eyers, P. A., Jones, A. R. & Eyers, C. E. (2017). 'Evaluation of Parameters for Confident Phosphorylation Site Localization Using an Orbitrap Fusion Tribrid Mass Spectrometer'. *J Proteome Res*. **16**. 3448-3459.